

Konsistenzprobleme der Data Envelopment Analysis in der empirischen Forschung

Dissertation der
Wirtschaftswissenschaftlichen Fakultät
der Universität Zürich
zur Erlangung der Doktorwürde eines
Doktors der Wirtschaftswissenschaften

vorgelegt von
Lukas Steinmann
von Zürich

genehmigt auf Antrag von
von Prof. Dr. Peter Zweifel

Die Wirtschaftswissenschaftliche Fakultät der Universität Zürich gestattet hierdurch die Drucklegung der vorliegenden Dissertation, ohne damit zu den darin ausgesprochenen Anschauungen Stellung zu nehmen.

Zürich, den 3. Juli 2002

Der Dekan: Prof. Dr. H. P. Wehrli

Vorwort

Die vorliegende Arbeit ist als Dissertationsschrift während meiner Zeit als wissenschaftlicher Mitarbeiter am Sozialökonomischen Institut der Universität Zürich entstanden. Ich bedanke mich an dieser Stelle herzlich bei all jenen, die mich in irgendeiner Form gefördert und unterstützt und somit diese Dissertation erst ermöglicht haben. Dieser Dank gilt insbesondere meiner Familie.

Ich danke Prof. Dr. Peter Zweifel in zweierlei Hinsicht. Erstens hat er mir im Herbst 1997 den Auftrag erteilt, eine Diplomarbeit „Zur Messung der Effizienz stationärer Leistungen“ zu verfassen und somit den Grundstein der vorliegenden Dissertation gelegt. Zweitens möchte ich Prof. Dr. Peter Zweifel dafür danken, dass er mir die grösstmöglichen Freiheiten gewährt und die Dissertation betreut hat.

Speziellen Dank gebührt meiner Ehefrau, Hanna, die stark an den Opportunitätskosten meiner Doktorandenzeit beteiligt war. Zudem danke ich den Mitarbeitern des Sozialökonomischen Instituts, ohne deren Mithilfe diese Arbeit nicht in dieser Form möglich gewesen wäre. Im speziellen danke ich Hansjörg Lehmann, Dr. Markus König, Yves Schneider (nicht nur für die inhaltlichen Anregungen, sondern auch für die vielen Tipps und Tricks zum Arbeiten mit \LaTeX), Harry Telser sowie meinem Bruder Christoph für die kritische Lektüre meiner Dissertation.

Zürich, im Juli 2002

Lukas Steinmann

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Einleitung und Fragestellung | 1 |
| 1.1 | Einleitung | 1 |
| 1.2 | Fragestellung | 4 |
| 1.3 | Gliederung | 5 |
| 2 | Die Effizienzmessung | 7 |
| 2.1 | Die Produktionstheorie | 7 |
| 2.1.1 | Die Darstellung der Technologiemenge | 7 |
| 2.1.2 | Die Axiome der Technologiemenge | 11 |
| 2.2 | Nichtparametrische Methoden der Effizienzmessung | 18 |
| 2.2.1 | Effizienz, Effizienzmasse und Effizienzwerte | 18 |
| 2.2.2 | Die Distanzfunktion | 19 |
| 2.2.3 | Die Data Envelopment Analysis (DEA) | 22 |
| 2.3 | Erweiterungen der DEA | 32 |
| 2.3.1 | Die One-Stage-Ansätze | 34 |
| 2.3.2 | Die Two-Stage-Ansätze | 35 |
| 3 | Konsistenzprobleme der Data Envelopment Analysis | 39 |
| 3.1 | Statistische Fundierung der DEA | 40 |
| 3.2 | Grafische Analyse | 43 |
| 3.2.1 | Die ideale Konstellation | 43 |
| 3.2.2 | Die nicht-ideale Konstellation | 46 |
| 3.3 | Mathematische Analyse | 52 |
| 3.3.1 | Mathematische Analyse der idealen Konstellation | 54 |
| 3.3.2 | Mathematische Analyse der nicht-idealen Konstellation | 57 |
| 3.3.3 | Mathematische Analyse der typischen Konstellation | 58 |

| | | |
|----------|---|------------|
| 3.3.4 | Fazit zur statistischen Fundierung der DEA | 62 |
| 3.4 | Konsistenzprobleme der DEA in der empirischen Forschung | 63 |
| 3.4.1 | Identifikation nicht-idealer Konstellationen | 63 |
| 3.4.2 | Eine Kennzahl zur Bewertung der Daten: Das Validitätsmass der Effizienzdifferenzen | 66 |
| 4 | Monte-Carlo-Simulation der Effizienzanalyse | 71 |
| 4.1 | Der Ablauf der Simulation | 71 |
| 4.2 | Phase 1: Der Datengenerierungsprozess | 74 |
| 4.3 | Phase 2: Berechnung der Daten | 77 |
| 4.4 | Phase 3: DEA, Analyse der Validität und Hypothesentests | 80 |
| 4.5 | Phase 4: Auswertung der Simulation | 83 |
| 4.5.1 | Geschätzte erwartete Verzerrungen | 83 |
| 4.5.2 | Hypothesentests | 89 |
| 4.6 | Verallgemeinerung der Resultate | 94 |
| 5 | Eine empirische Anwendung der DEA | 97 |
| 5.1 | DEA schweizerischer Krankenhäuser | 97 |
| 5.1.1 | Die Inputs und Outputs der Krankenhäuser | 98 |
| 5.1.2 | Die Daten der Stichprobe | 100 |
| 5.1.3 | Die Restriktionen der Grenzzraten der Substitution und Transfor- mation | 100 |
| 5.1.4 | Die Ergebnisse der DEA | 102 |
| 5.1.5 | Die ökonometrische Auswertung der DEA-Ergebnisse | 103 |
| 5.1.6 | Bewertung der Schätz-Resultate | 110 |
| 5.2 | Bewertung der möglichen Ursachen der schwachen Erklärungskraft | 111 |
| 5.2.1 | Verzernte Effizienzdifferenzen als mögliche Ursache | 111 |
| 5.2.2 | Unterschiede in der Qualität und im Case-Mix als mögliche Ursache | 118 |
| 5.3 | Fazit zur DEA im schweizerischen Krankenhaussektor | 119 |
| 6 | Schlussfolgerungen | 123 |
| A | MATLAB Monte-Carlo-Syntax | 129 |
| B | Fehlerwahrscheinlichkeiten | 133 |

Abbildungsverzeichnis

| | | |
|------|--|----|
| 2.1 | Das Inputset | 9 |
| 2.2 | Das Outputset | 10 |
| 2.3 | Dominierende und dominierte Input- und Outputvektoren | 12 |
| 2.4 | Konvexe Input- und Outputsets | 13 |
| 2.5 | Konvexe bzw. nicht-konvexe Technologiemenzen | 14 |
| 2.6 | Kostenlose Entsorgung von Inputs und Outputs | 15 |
| 2.7 | Die Art der Skalenerträge | 16 |
| 2.8 | Die Technologiemenze | 22 |
| 2.9 | DEA mit variablen Skalenerträgen | 28 |
| 2.10 | Linked-Ratio-Modelle | 31 |
| 3.1 | Die ideale Konstellation mit unterschiedlich effizienten Gruppen | 44 |
| 3.2 | Die ideale Konstellation mit gleich effizienten Gruppen | 45 |
| 3.3 | Die nicht-ideale Konstellation | 47 |
| 3.4 | Gruppenspezifische Ineffizienz | 49 |
| 3.5 | Klassifizierung der Abschnitte | 51 |
| 3.6 | Mit der Gruppenzugehörigkeit korrelierte Inputverhältnisse | 65 |
| 3.7 | Die beiden Komponenten des Validitätsmasses | 68 |
| 4.1 | Flussdiagramm der Monte-Carlo-Simulation | 73 |
| 4.2 | Illustration des Verschiebungsparameters ι | 76 |

Tabellenverzeichnis

| | | |
|------|---|-----|
| 4.1 | Parameterwerte der Simulation | 78 |
| 4.2 | Erwartete Fehlerwahrscheinlichkeiten der Hypothesentests | 82 |
| 4.3 | Geschätzte erwartete Verzerrungen der Gruppe A, Teil 1 | 84 |
| 4.4 | Geschätzte erwartete Verzerrungen der Gruppe A, Teil 2 | 85 |
| 4.5 | Geschätzte erwartete Verzerrungen der Gruppe B | 86 |
| 4.6 | Die geschätzten erwarteten Verzerrungen der Effizienzdifferenzen | 87 |
| 4.7 | Theoretische Validität der DEA-Effizienz | 89 |
| 4.8 | Anteil α -Fehler ($\iota^A = 0$) | 90 |
| 4.9 | Anteil β -Fehler ($\iota^A = 0.05$) | 92 |
| 4.10 | Anteil β -Fehler ($\iota^A = 0.10$) | 92 |
| 4.11 | Anteil β -Fehler ($\iota^A = 0.15$) | 93 |
| 4.12 | Anteil β -Fehler ($\iota^A = 0.20$) | 94 |
| 5.1 | Unterstellte Wertgrenzen der Produktionsfaktoren | 101 |
| 5.2 | Resultate der Random-Effects-Schätzung | 108 |
| 5.3 | Gruppenspezifische Abschnitte: Rechtsträger | 112 |
| 5.4 | Gruppenspezifische Abschnitte: Subventionierung | 113 |
| 5.5 | Gruppenspezifische Abschnitte: Defizitdeckung | 114 |
| 5.6 | Gruppenspezifische Abschnitte: Privatspital | 114 |
| 5.7 | Gruppenspezifische Abschnitte: Pauschalsubvention | 115 |
| 5.8 | Validitätsmasse der Effizienzdifferenzen in Bezug auf die Anreizvariablen | 116 |
| B.1 | Fehlerwahrscheinlichkeiten | 133 |

Kapitel 1

Einleitung und Fragestellung

1.1 Einleitung

Der ökonomische Effizienzbegriff geht auf Pareto (1897) zurück. Ein Zustand ist paretoeffizient, wenn es nicht möglich ist, ein Individuum besser zu stellen ohne dadurch andere schlechter zu stellen. Dieser Effizienzbegriff wurde von Koopmans (1951) auf die Produktionstheorie übertragen. Ein Herstellungsprozess ist dann effizient, wenn es unmöglich ist, mehr von einem Gut herzustellen ohne gleichzeitig mehr Produktionsfaktoren einzusetzen und es auch nicht möglich ist, weniger Produktionsfaktoren einzusetzen ohne einen Rückgang bei der Produktion hinzunehmen. Bezieht sich dieser Effizienzbegriff ausschliesslich auf Mengengrössen, beschreibt er die technische Transformation von Produktionsfaktoren in Güter. Ein Herstellungsprozess ist *technisch ineffizient*, wenn eine gegebene Menge Güter auch mit kleineren Mengen von Produktionsfaktoren produziert werden kann bzw. wenn mit einer vorgegebenen Menge an Produktionsfaktoren mehr Güter hergestellt werden können.

Aus ökonomischer Sicht sind neben der technischen Effizienz weitere Effizienzkonzepte von Bedeutung. Ein Herstellungsprozess ist *kosteneffizient*, wenn ein Güterbündel mit minimalen Kosten produziert wird. Die Kosteneffizienz bezieht sich nur noch bei den hergestellten Gütern auf Mengengrössen; bei den Produktionsfaktoren ist nun *eine* Wertgrösse – die zu Marktpreisen bewerteten Produktionsfaktoren – relevant. Andererseits ist ein Herstellungsprozess *erlöseffizient*, wenn mit gegebenen Produktionsfaktoren ein Güterbündel produziert wird, das den maximalen Erlös generiert. Dieser Effizienzbegriff beinhaltet Mengengrössen bei den Produktionsfaktoren und *eine* Wertgrösse, den Erlös.

Die technische Effizienz ist eine notwendige Bedingung für die Kosten- und Erlöseffizienz.

Bei diesen Effizienzkonzepten waren jeweils die Produktionsfaktoren oder das zu produzierende Güterbündel vorgegeben. Ein weiteres Kriterium ist die Wahl der optimalen Menge des Güterbündels: Ein Herstellungsprozess ist *skaleneffizient*, wenn eine prozentuale Ausdehnung aller Produktionsfaktoren zu einer entsprechenden prozentualen Zunahme des Güterbündels führt. So lange eine Ausdehnung der Produktionsfaktoren zu einer überproportionalen Zunahme des Güterbündels führt, kann durch eine Ausweitung der Produktion das Wertverhältnis zwischen dem Güterbündel und den Produktionsfaktoren – unabhängig von den Preisen – gesteigert werden. Führt dagegen eine Ausdehnung der Produktionsfaktoren zu einer unterproportionalen Zunahme des Güterbündels, kann dieses Verhältnis durch eine *Reduktion* aller Produktionsfaktoren erhöht werden. Zur Erreichung der ökonomischen Effizienz ist es notwendig, dass ein Herstellungsprozess *kosten-, erlös- und skaleneffizient* ist. Für das ökonomische Effizienzkriterium sind ausschliesslich Wertgrössen relevant: Die ökonomische Effizienz kann in Geldeinheiten gemessen werden.

In einem Wettbewerbsmarkt gilt die Null-Gewinn-Bedingung. Diese Bedingung besagt, dass zusätzliche Anbieter in den Markt eintreten, bis keine Gewinne mehr erzielt werden. In einem Wettbewerbsmarkt kann ein Hersteller folglich im besten Fall – d.h. wenn er ökonomisch effizient produziert – alle Kosten decken. Ineffiziente Firmen realisieren Verluste und verschwinden vom Markt. Folglich werden in einem Wettbewerbsmarkt nur die effizienten Hersteller überleben. Kann sichergestellt werden, dass die Marktkräfte wirken, erübrigt sich die Effizienzmessung; Ressourcen dafür einzusetzen wäre Verschwendung, da die Marktkräfte die Ineffizienz beseitigen.

In modernen Volkswirtschaften wirken die Marktkräfte in verschiedenen Bereichen nicht oder nur teilweise bzw. sie werden bewusst eingeschränkt:

1. Bei *natürlichen Monopolen*. Natürliche Monopole sind durch abnehmende Grenzkosten gekennzeichnet. Aus wohlfahrtsökonomischen Überlegungen ist es effizient, wenn ein einzelner Anbieter die Versorgung sicherstellt. Zwei Probleme ergeben sich bei natürlichen Monopolen. Einerseits existiert für das entsprechende Gut kein Marktpreis, der den Monopolisten zur Effizienz zwingt. Andererseits ist der Monopolist bestrebt, ein aus wohlfahrtsökonomischen Gesichtspunkten zu geringes Versorgungsniveau anzustreben. Als Beispiele sind die Telekommunikationsbranche und die Stromverteilung zu nennen.
2. Beim Vorliegen *asymmetrischer Information* kann es ebenfalls zu einem Marktver-

sagen kommen. Die asymmetrische Information führt dazu, dass Güter infolge von *Adverse Selection* nicht angeboten werden (z.B. im Versicherungssektor) oder dass diese es den Akteuren erlaubt, eine Informationsrente zu erzielen (*Moral Hazard*, z.B. im Gesundheitssektor). Typischerweise führen die negative Auslese und das *Moral Hazard* zu einem niedrigeren Wohlfahrtsniveau.

3. Bei *öffentlichen Gütern* kommt es zu einem Marktversagen, da diese Güter nicht bzw. in einem suboptimalen Ausmass bereitgestellt (*Free-Rider*-Problematik) oder mangels geregelter Eigentumsrechte von den Akteuren zu intensiv genutzt werden (externe Effekte). Öffentliche Güter umfassen z.B. das Rechtssystem, die Landesverteidigung und den Umweltschutz.
4. Falls *Verteilungsziele* im Vergleich zu den Allokationszielen politisch eine übergeordnete Rolle einnehmen, werden die Marktkräfte häufig bewusst eingeschränkt. Als Beispiel ist der Bildungssektor zu nennen, der grundsätzlich allen offen stehen soll mit dem Ziel, die Startchancen auszugleichen.

In diesen Bereichen scheint es sinnvoll zu sein, dass der Staat regulierend eingreift, um die fehlenden Marktkräfte durch andere Mechanismen zu ersetzen oder die entsprechenden Aufgaben an Institutionen überträgt. Allerdings stehen den potentiellen Wohlfahrtsge winnen auch Kosten gegenüber. Die Neue Theorie der Regulierung befasst sich mit der Fragestellung, wie die Akteure dazu angehalten werden können, die übergeordneten Ziele auf effiziente Weise zu verfolgen [vgl. Laffont und Tirole (1993)]. Als Kern des Problems erweist sich dabei die asymmetrische Information zwischen den Akteuren und den Regulierungsbehörden. Die Akteure haben einen Informationsvorsprung, den sie dazu nutzen können, eine Informationsrente zu erzielen. Teil des Informationsvorsprungs der Akteure ist die Kenntnis der eigenen Effizienz.¹ Informationen bezüglich der Effizienz, die das Erzielen von Renten verhindern, haben folglich für die Regulierungsbehörden einen Wert. Laffont und Tirole (1993) sind der Ansicht (S. 3):

The presence of moral hazard and adverse selection, and the concomitant loss of control for the regulator, create a demand for information gathering.

¹In der Neuen Theorie der Regulierung wird der Informationsvorsprung in zwei Komponenten unterteilt. Einerseits verfügt der Akteur über Informationen bezüglich des erbrachten *Efforts* (*hidden action*) und andererseits bezüglich der eingesetzten *Technologie* (*hidden information*). Im Verbund können diese Begriffe auch als Effizienz interpretiert werden.

Die Effizienzmessung kann somit helfen, diesen Informationsvorsprung zu reduzieren und die Informationsrente der Akteure zu senken. Aus ökonomischer Sicht ist es daher sinnvoll, in bestimmten Bereichen der Wirtschaft Effizienzmessungen durchzuführen.

1.2 Fragestellung

Im Zentrum der vorliegenden Arbeit steht die Frage, ob und unter welchen Umständen die Effizienzmessung mittels der *Data Envelopment Analysis* (DEA) konsistent erfolgen kann. Dabei wird unter dem Begriff der Konsistenz ein statistisches Qualitätsmerkmal verstanden, das eine Voraussetzung für den fehlerfreien Rückschluss von einer Stichprobe auf die Grundgesamtheit darstellt. Es stellt sich die Frage, unter welchen Umständen diese Methode in der empirischen Forschung zu korrekten Effizienzaussagen führt.

Die DEA erfreut sich seit ihrer Entwicklung durch Charnes, Cooper und Rhodes (1978) wachsender Beliebtheit in der sozial- und naturwissenschaftlichen Forschung. Die wenig restriktiven Annahmen, die bescheidenen Anforderungen an die Daten² und die Möglichkeit, komplexe Transformationsprozesse zu bewerten, haben der DEA auch in der volkswirtschaftlichen Forschung – insbesondere bei der Analyse regulierter Märkte – eine bedeutende Rolle zukommen lassen.³

Anhand eines Beispiels soll gezeigt werden, wieso die Konsistenz der Methode der Effizienzmessung im Bereich der Regulierung einen besonderen Stellenwert einnimmt: Wird in einem Konzern eine Filiale aufgrund ihrer vermeintlich hohen Effizienz, die aber fehlerhaft ermittelt wurde, zu Lasten einer an sich effizienteren Filiale nicht geschlossen, mögen die Konsequenzen zwar kostspielig, aus volkswirtschaftlicher Perspektive aber kaum gravierend sein.⁴ Wenn dagegen die Massnahmen zur Beseitigung der Wohlfahrtsverluste in einem regulierten Markt aufgrund von Fehlentscheidungen nicht greifen, entsprechen die resultierenden Kosten dem diskontierten Wohlfahrtsverlust. Da zusätzlich noch die Kosten der Regulierung⁵ anfallen, wird insgesamt ein im Vergleich zur nicht regulierten

²Informationen über die Preise der Produktionsfaktoren und Produkte sind nicht notwendig.

³Die DEA wurde vor allem zur Effizienzmessung im Gesundheits-, Bildungs- Transport- und Stromverteilungssektor eingesetzt.

⁴Ein typisches Anwendungsgebiet der DEA ist das *Benchmarking*, das in den Bereich des *Controlling* gehört.

⁵Die Kosten der Regulierung umfassen einerseits die Opportunitätskosten der Ressourcen, die für die Regulierung eingesetzt werden, und andererseits die gesamtwirtschaftlichen Effizienzverluste, die bei der Finanzierung der Regulierung anfallen (verzerrende Wirkung des Steuersystems).

Ausgangslage inferiorer Zustand realisiert. Wird den Akteuren durch die Fehlentscheidungen zusätzlicher Spielraum zur Erzielung von Renten geschaffen, können die Kosten noch höher ausfallen. Verstärkt wird die Problematik durch die Tatsache, dass Anpassungen der Regulierung mit Transaktionskosten verbunden sind, da die Rentenempfänger sich gegen solche Veränderungen wehren werden. Zudem können Anpassungen nicht ohne zeitliche Verzögerung realisiert werden, was die umgehende Beseitigung des inferioren *Status quo* verhindert. Daher ist es von herausragender Bedeutung, dass die Methoden, die als Grundlage der Regulierung dienen, nicht zu falschen Schlussfolgerungen verleiten.

1.3 Gliederung

In Kapitel 2 erfolgt zuerst eine Einführung in die Produktionstheorie, welche analysiert, in welchem Ausmass aus Produktionsfaktoren – wie z.B. Arbeit und Kapital – Zwischen- und Endprodukte hergestellt werden können. Anschliessend wird der erste (nichtparametrische) Ansatz zur Messung der Effizienz hergeleitet, der zur Entwicklung der DEA geführt hat. Die DEA, als Forschungsgegenstand dieser Arbeit, wird in einem eigenen Abschnitt hergeleitet. Der letzte Abschnitt befasst sich mit der Frage, wie heterogene Rahmenbedingungen in die Effizienzmessung integriert werden sollen. Es besteht ein Konsens, dass unterschiedliche Rahmenbedingungen berücksichtigt werden müssen – die konkrete Integration dieser Rahmenbedingungen ist dabei aber kontrovers.

In Kapitel 3 wird anschliessend untersucht, unter welchen Umständen die Resultate der DEA für eine statistische Analyse geeignet sind. Im Zentrum der Fragestellung stehen die statistischen Eigenschaften der Schätzer der Effizienz. Die Voraussetzungen für eine statistische Analyse der DEA-Resultate sind an verschiedene Verteilungsannahmen geknüpft, welche sich als derart restriktiv erweisen, dass in der empirischen Forschung nicht *a priori* davon ausgegangen werden darf, dass diese erfüllt sind. Zudem lässt sich in der empirischen Forschung nicht überprüfen, ob diese Annahmen im Einzelfall erfüllt sind. Zur Lösung des Problems wird eine Kennzahl konstruiert, die eine Bewertung der Daten und einen Rückschluss auf die Validität der verwendeten Methoden zulässt.

In Kapitel 4 werden die Hypothesen des dritten Kapitels anhand einer *Monte-Carlo*-Simulation überprüft. Bei der *Monte-Carlo*-Simulation wird ein „wahres“ Modell definiert und anschliessend werden unter Zuhilfenahme von Zufallszahlen (mit bekannter Verteilung) künstliche Daten generiert. Die künstlichen Daten werden dann in einer DEA ausgewertet. Da das „wahre“ Modell, die Zufallszahlen bzw. deren Verteilung und folglich

die „wahre“ Effizienz bekannt sind, ist der Rückschluss auf allfällige durch die DEA gemachten Fehler zulässig. Es werden verschiedene „wahre“ Modelle überprüft und es wird gezeigt, welche Eigenschaften des „wahren“ Modells zu Fehlern (d.h. zu Verzerrungen) bei der DEA führen. Schliesslich wird untersucht, wie sich Verzerrungen bei der DEA auf die Stufe der statistischen Auswertung der DEA-Resultate übertragen.

In Kapitel 5 wird eine DEA mit ökonometrischer Auswertung für den schweizerischen Krankenhaussektor präsentiert. Es zeigt sich, dass die geschätzte Effizienz kaum in einem statistisch signifikanten Zusammenhang mit den Rahmenbedingungen steht. Verschiedene Ursachen für den schwachen Erklärungsgehalt der Auswertung der DEA-Resultate werden diskutiert. Im Vordergrund steht aber die Frage, ob die schwache Erklärungskraft auf verzerrte DEA-Resultate zurückzuführen ist.

In Kapitel 6 werden die Resultate der Arbeit zusammengefasst, diskutiert und zukünftige Forschungsschwerpunkte vorgeschlagen.

Kapitel 2

Die Effizienzmessung

In diesem Kapitel werden die theoretischen Grundlagen für die folgenden Ausführungen erarbeitet. Zuerst folgt in Abschnitt 2.1 eine Einführung in die Produktionstheorie; die Technologiemenge, d.h. die Menge der Produktionsmöglichkeiten und deren Eigenschaften werden definiert. Die Distanzfunktion – der erste Ansatz der nichtparametrischen Effizienzmessung – wird anschliessend direkt aus der Darstellung der Technologiemenge abgeleitet. In Abschnitt 2.2 wird die *Data Envelopment Analysis* (DEA), eine Weiterentwicklung der Distanzfunktion, ausführlich präsentiert. Im Abschnitt 2.3 werden Erweiterungen der DEA diskutiert, die die Berücksichtigung heterogener Rahmenbedingungen zulassen.

2.1 Die Produktionstheorie

2.1.1 Die Darstellung der Technologiemenge

Die in diesem Abschnitt präsentierte Definition der Technologiemenge geht auf Koopmans (1951) und Debreu (1951) zurück und wurde im Rahmen der *Activity Analysis* eingeführt. Bei der Technologiemenge handelt es sich in diesem Zusammenhang um eine sehr allgemein gehaltene Aussage darüber, welche Güter und Dienstleistungen mit welchen Produktionsfaktoren produziert werden können. Der technische Herstellungsprozess an sich steht dabei nicht im Vordergrund; im Gegenteil, er wird in der Regel vollständig ausgeklammert und als *Black Box* betrachtet. Die Begründung für diese abstrahierende Betrachtungsweise ist, dass die Effizienzmessung nicht ein Leistungs- oder Produktivitätsvergleich zwischen verschiedenen Maschinen oder unterschiedlich ausgebildeten Ar-

beitskräften darstellt. Viel mehr steht der gesamte Produktionsprozess im Zentrum der Fragestellung. Die Technologiemenge ist durch eine Reihe verschiedener Inputs zur Produktion von Outputs charakterisiert. Sowohl die m Inputs $X = (x_1, \dots, x_m)$, $X \in \mathbb{R}_+^m$ wie auch die s Outputs $Y = (y_1, \dots, y_s)$, $Y \in \mathbb{R}_+^s$ sind (nicht-negative) Mengeneinheiten, die während einer bestimmten Zeit in die Produktion einfließen bzw. als Ergebnis dieses Prozesses als Güter und Dienstleistungen für den Konsum oder zur weiteren Verarbeitung zur Verfügung stehen.

Die Gesamtheit aller zulässigen Input-Outputtransformationen (jede einzelne stellt eine Technologie oder – wie in der *Activity Analysis* bezeichnet – eine Aktivität dar) wird als Technologiemenge bezeichnet und ist definiert durch

$$\mathcal{T} = \{(Y, X) \in \mathbb{R}_+^{s+m} \mid Y \text{ kann mit } X \text{ produziert werden}\}. \quad (2.1)$$

Die Technologiemenge widerspiegelt die technischen Produktionsmöglichkeiten, d.h. die Möglichkeit, Inputs in Outputs zu transformieren: Mit einer gegebenen Menge an Inputs kann in einem Zeitraum eine bestimmte Menge Outputs produziert werden. Diese triviale Aussage kann einerseits empirisch oder theoretisch erfolgen. In der empirischen Variante – dem *nichtparametrischen* Ansatz – definieren realisierte, d.h. beobachtete Input-Outputtransformationen die Technologiemenge. Die Legitimation der Technologiemenge ist garantiert, da die beobachteten Input- und Outputmengen effektiv absorbiert bzw. realisiert wurden. In der theoretischen Variante wird eine Produktionsfunktion unterstellt und der Anspruch erhoben, dass ein funktionaler Zusammenhang zwischen Inputs und Outputs besteht. Diese Variante wird *parametrischer* Ansatz genannt. In den folgenden Ausführungen wird das Schwergewicht auf die nichtparametrische Methode gelegt.¹ Die beiden Methoden haben aber vieles gemein und es wird – wo im Gesamtzusammenhang von Bedeutung – die Verbindung zu den parametrischen Methoden bewusst hergestellt.

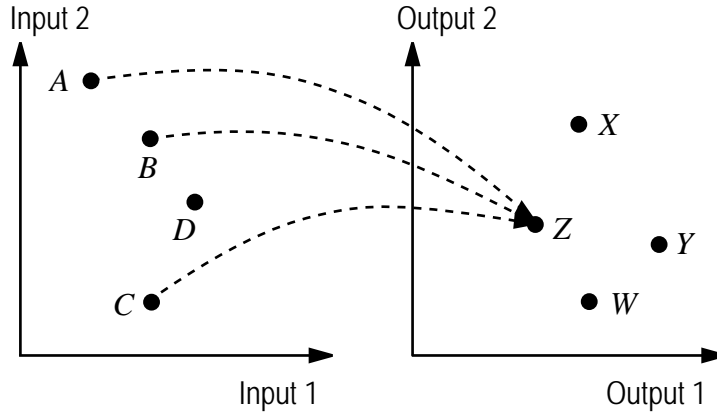
Aus der Technologiemenge lässt sich für jeden Outputvektor $Y \in \mathcal{T}$ ein Inputset $\mathcal{L}(Y)$ bestimmen

$$\mathcal{L}(Y) = \{X \in \mathbb{R}_+^m \mid (Y, X) \in \mathcal{T}\}. \quad (2.2)$$

Jedes Inputset umfasst alle Inputvektoren, mit denen der betrachtete Outputvektor Y hergestellt werden kann. Das Inputset widerspiegelt somit die Substitutionsbeziehungen zwischen den einzelnen Inputs. In Abbildung 2.1 ist eine Situation abgebildet, in der zwei

¹Eine Übersicht zu den parametrischen Methoden der Effizienzmessung bieten z.B. Kumbhakar und Lovell (2000) oder Coelli, Rao und Battese (1998).

Inputs in zwei verschiedene Outputs transformiert werden können. Der Outputvektor Z kann mit jedem der drei Inputvektoren A , B und C produziert werden. Das Inputset $\mathcal{L}(Z)$ beinhaltet folglich die Inputvektoren A , B und C .

Abbildung 2.1: Das Inputset $\mathcal{L}(Z)$ 

Analog erfolgt die Definition der Outputsets

$$\mathcal{P}(X) = \{Y \in \mathbb{R}_+^s \mid (Y, X) \in \mathcal{T}\}. \quad (2.3)$$

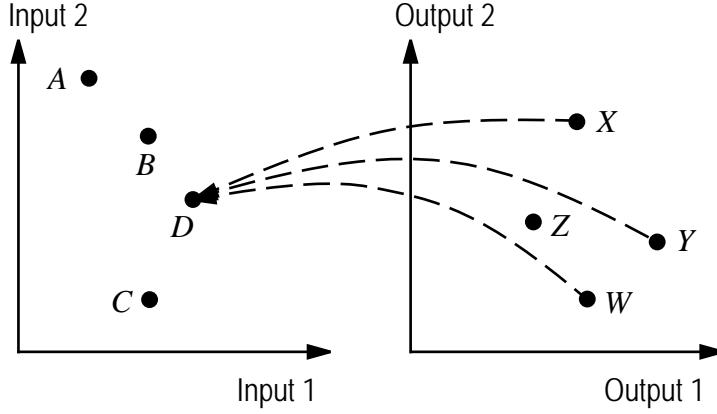
Jedes Outputset gibt an, welche Outputvektoren mit einem gegebenen Inputvektor produziert werden können und unterstreicht im Gegensatz zum Inputset die Substitutionsbeziehungen zwischen den Outputs. Das Outputset $\mathcal{P}(D)$ umfasst gemäss Abbildung 2.2 die Outputvektoren W , X und Y . Aus den beiden Inputmengen, die durch den Inputvektor D beschrieben werden, lassen sich folglich entweder die beiden Outputmengen gemäss Outputvektor W , X oder Y produzieren.²

Die inverse Beziehung zwischen Input- und Outputsets lässt sich anhand der Technologiemenge \mathcal{T} einfach nachweisen. Aus Gleichung 2.2 $X \in \mathcal{L}(Y) \Leftrightarrow (X, Y) \in \mathcal{T}$ und Gleichung 2.3 $Y \in \mathcal{P}(X) \Leftrightarrow (X, Y) \in \mathcal{T}$ folgt

$$X \in \mathcal{L}(Y) \Leftrightarrow Y \in \mathcal{P}(X). \quad (2.4)$$

Ein Outputvektor Y lässt sich nur dann aus einem bestimmten Inputvektor X herstellen, wenn mit demselben Inputvektor X mindestens der Outputvektor Y hergestellt werden

²Im Produktionsprozess zeigen die Pfeile natürlich in die entgegengesetzte Richtung. Die mathematische Abbildung erfolgt aber wie in Abbildung 2.2.

Abbildung 2.2: Das Outputset $\mathcal{P}(D)$ 

kann. Die Technologiemenge lässt sich somit auch durch die Inputsets und die Outputsets darstellen

$$\begin{aligned}\mathcal{T}(\mathcal{L}(), \mathcal{P}()) &\equiv \{(Y, X) \in \mathbb{R}_+^{s+m} \mid X \in \mathcal{L}(Y)\} \\ &\equiv \{(Y, X) \in \mathbb{R}_+^{s+m} \mid Y \in \mathcal{P}(X)\}.\end{aligned}\quad (2.5)$$

Eine alternative Darstellungsform der Technologie, die sich im Verlauf dieses Kapitels bewähren wird, ist in Gleichung 2.6 aufgeführt:

$$\mathcal{T}(\mathcal{L}(), \mathcal{P}()) \equiv \{(Y, X) \in \mathbb{R}_+^{s+m} \mid F(X, Y) \leq 0\}.\quad (2.6)$$

In Gleichung 2.6 wurde die Bedingung, dass die Inputs im Inputset bzw. die Outputs im Outputset enthalten sein müssen, durch die Bedingung, dass die Transformationsfunktion $F(X, Y)$ kleiner gleich null sein muss, ersetzt [vgl. z.B. Varian (1995), Kapitel 1 oder Mas-Colell, Whinston und Green (1995), Kapitel 5]. Ohne weitere Spezifikation dieser Funktion kann allgemein gesagt werden, dass die aggregierten Outputs minus die aggregierten Inputs kleiner gleich null sein müssen. Werden zusätzliche Inputs zur Produktion eines gegebenen Outputs absorbiert, sinkt *ceteris paribus* der Wert der Transformationsfunktion. War der Funktionswert in der Ausgangslage kleiner gleich null, muss er auch nachher noch kleiner gleich null sein. Werden die Inputs aber bei gegebenem Output reduziert, steigt der Funktionswert. Nimmt die Transformationsfunktion schliesslich den Wert null an, lassen sich die Inputs bei gegebenem Output nicht weiter reduzieren – die Grenze

der Produktionsmöglichkeiten (GPM) ist dann erreicht. Die Bedingung $F(X, Y) \leq 0$ garantiert, dass die Input-Outputtransformationen Bestandteil der Technologiemenge sind.

2.1.2 Die Axiome der Technologiemenge

Mit der oben beschriebenen Technologiemenge lassen sich erst schwache Aussagen über die technische Effizienz von Input-Outputtransformationen ableiten. Alle diesbezüglichen Aussagen beschränken sich auf den Vergleich der Inputvektoren jeweils eines Inputsets; d.h. es können nur die Inputvektoren, die die Produktion eines spezifischen Outputvektors ermöglichen, miteinander verglichen werden. Inputvektoren, die verschiedenen Inputsets angehören, können nicht miteinander verglichen werden. Dasselbe gilt für die Outputvektoren: Es können nur Outputvektoren miteinander verglichen werden, die demselben Outputset angehören.

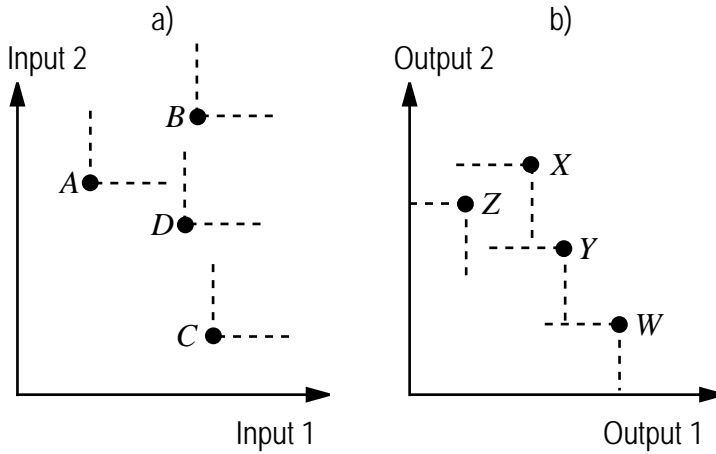
Zudem ist es bei der Betrachtung eines einzelnen Input- oder Outputsets nicht möglich, das Ausmass der Effizienz der verschiedenen Input- bzw. Outputvektoren zu bestimmen. In diesem Stadium ist erst die Unterscheidung zwischen dominierten und dominierenden Input-Outputtransformationen zulässig. Zur Illustration ist in Abbildung 2.3a) das Inputset *eines* Outputvektors dargestellt. Die Inputvektoren A und D dominieren den Inputvektor B strikt, da beide sowohl von Input 1 und 2 kleinere Mengen zur Herstellung des Outputvektors benötigen. Der Inputvektor C dominiert in diesem Beispiel dagegen keine Inputvektoren, wird aber auch selbst nicht dominiert. Über die Effizienz des Inputvektors B kann keine Aussage gemacht werden, da die Isoquante und insbesondere die effizienten Abschnitte der Isoquante noch nicht definiert sind. In Abbildung 2.3b) ist eine vergleichbare Situation anhand eines Outputsets dargestellt. Der Outputvektor X dominiert Z strikt; Y und W dominieren keinen Outputvektor, werden selbst aber auch nicht dominiert.

Erst zusätzliche Annahmen – die folgenden Axiome – bezüglich der Eigenschaften der Input- und Outputisoquanten und der Beziehung zwischen den einzelnen Isoquanten definieren die Struktur, d.h. die geometrischen Eigenschaften der Technologiemenge:

Axiom 1 Konvexe Inputsets: Wenn $(X_j) \in \mathcal{L}(Y)$, $j = 1, \dots, n$ und $\lambda_j \geq 0$, $j = 1, \dots, n$ und $\sum_{j=1}^n \lambda_j = 1$, dann ist auch $\sum_{j=1}^n \lambda_j X_j \in \mathcal{L}(Y)$

Axiom 2 Konvexe Outputsets: Wenn $(Y_j) \in \mathcal{P}(X)$, $j = 1, \dots, n$ und $\lambda_j \geq 0$, $j = 1, \dots, n$ und $\sum_{j=1}^n \lambda_j = 1$, dann ist auch $\sum_{j=1}^n \lambda_j Y_j \in \mathcal{P}(X)$

Abbildung 2.3: Dominierende und dominierte Input- und Outputvektoren



Axiom 3 Konvexe Technologie: Wenn $(X_j, Y_j) \in \mathcal{T}$, $j = 1, \dots, n$ und $\lambda_j \geq 0$, $j = 1, \dots, n$ und $\sum_{j=1}^n \lambda_j = 1$, dann ist auch $\sum_{j=1}^n \lambda_j (X_j, Y_j) \in \mathcal{T}$

Axiom 4 Kostenlose Entsorgung von Inputs:

- a) in der starken Form verlangt, dass wenn $(X, Y) \in \mathcal{T}$ und $\bar{X} \geq X$, dann ist auch $(\bar{X}, Y) \in \mathcal{T}$
- b) in der schwachen Form verlangt, dass wenn $(X, Y) \in \mathcal{T}$, dann ist auch $(\theta X, Y) \in \mathcal{T} \forall \theta \geq 1$.

Axiom 5 Kostenlose Entsorgung von Outputs:

- a) in der starken Form verlangt, dass wenn $(X, Y) \in \mathcal{T}$ und $0 \leq \bar{Y} \leq Y$, dann ist auch $(X, \bar{Y}) \in \mathcal{T}$
- b) in der schwachen Form verlangt, dass wenn $(X, Y) \in \mathcal{T}$, dann ist auch $(X, \theta Y) \in \mathcal{T} \forall 0 \leq \theta \leq 1$.

Axiom 6 Die Art der Skalenerträge:

- a) Konstante Skalenerträge (*Constant Returns to Scale, CRS*): Wenn $(X, Y) \in \mathcal{T}^{CRS}$, dann ist $(\kappa X, \kappa Y) \in \mathcal{T}^{CRS}$, $\forall \kappa \geq 0$
- b) Nichtzunehmende Skalenerträge (*Non Increasing Returns to Scale, NIR*): Wenn $(X, Y) \in \mathcal{T}^{NIR}$, dann ist $(\kappa X, \kappa Y) \in \mathcal{T}^{NIR}$, $\forall 0 \leq \kappa \leq 1$
- c) Nichtabnehmende Skalenerträge (*Non Decreasing Returns to Scale, NDR*): Wenn

$(X, Y) \in \mathcal{T}^{NDR}$, dann ist $(\kappa X, \kappa Y) \in \mathcal{T}^{NDR}$, $\forall \kappa \geq 1$)

d) Variable Skalenerträge (*Variable Returns to Scale*, *VRS*) liegen dann vor, wenn weder Axiom 6a), b) oder c) zutreffen. Die Technologiemenge mit variablen Skalenerträgen ist die Schnittmenge der beiden Technologiemengen mit nichtzu- und nichtabnehmenden Skalenerträgen: $\mathcal{T}^{VRS} = \mathcal{T}^{NDR} \cap \mathcal{T}^{NIR}$.

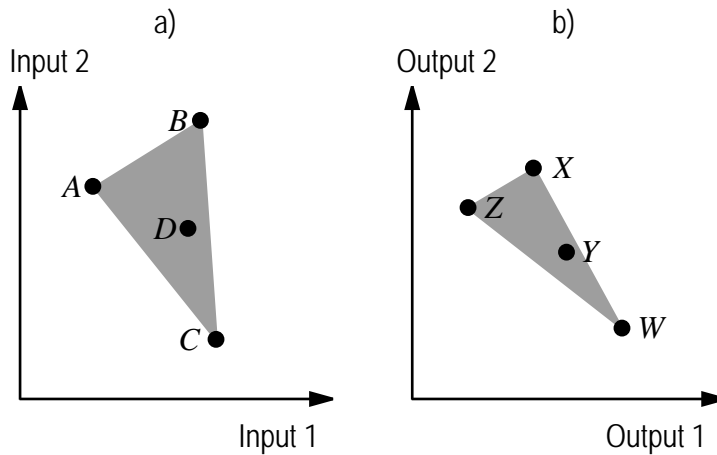
Axiom 7 Empirische Vollständigkeit: Eine Technologiemenge ist empirisch vollständig, wenn alle empirisch beobachteten Input-Outputkombinationen in der Technologiemenge enthalten sind: $(X_j, Y_j) \in \mathcal{T}$, für $j = 1, \dots, n$.

Axiom 8 Minimale Extrapolation: \mathcal{T}' ist die kleinstmögliche Teilmenge in \mathcal{T} , die den Axiomen 1, 2 und 4 bis 7 oder 3 bis 7 genügt: $\mathcal{T}' \subset \mathcal{T}$.

Das erste Axiom besagt, dass die konvexen Linearkombinationen von Inputvektoren eines Inputsets ebenfalls zum Inputset gehören. Jeder Punkt auf der grau schattierten Fläche in Abbildung 2.4a) gehört folglich ebenfalls zum Inputset $\mathcal{L}(Y) = (A, B, C, D)$.

Das zweite Axiom verlangt, dass die konvexen Linearkombinationen von Outputvektoren eines Outputsets ebenfalls zum Outputset gehören; somit gehören alle Punkte auf der grauen Fläche in Abbildung 2.4b) zum Outputset.

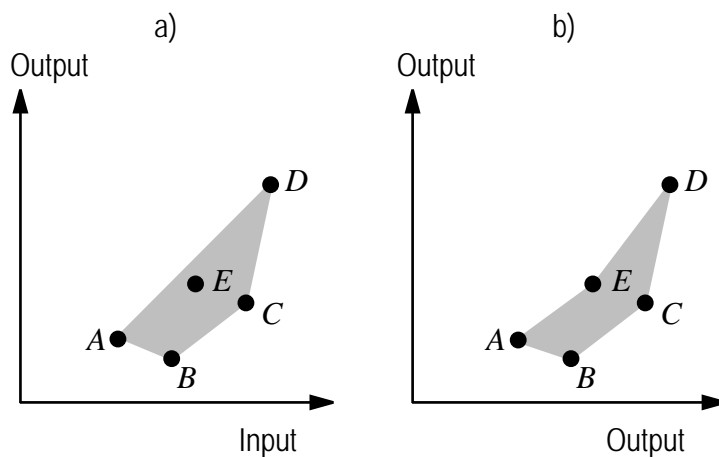
Abbildung 2.4: Konvexe Input- und Outputsets



Das dritte Axiom impliziert, dass konvexe Linearkombinationen von Input-Outputtransformationen der Technologiemenge ebenfalls zur Technologiemenge gehören. Dieses

Axiom gewährleistet, dass die beiden ersten Axiome erfüllt sind. Die beiden ersten Axiome implizieren aber nicht das dritte Axiom. Zur Illustration dient Abbildung 2.5. Bei beiden Technologiemengen seien die Input- und Outputsets konvex. In der Abbildung 2.5a) ist eine konvexe und in Abbildung 2.5b) eine nicht-konvexe Technologiemenge dargestellt. Ist das dritte Axiom erfüllt bzw. wird dieses unterstellt, lassen sich Technologiemengen mit zunehmender Skalenelastizität, d.h. global zunehmenden Skalenerträgen, nicht adäquat darstellen.

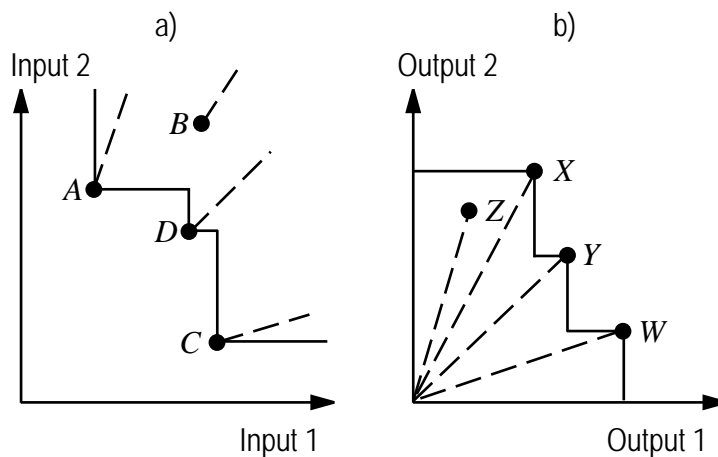
Abbildung 2.5: Konvexe bzw. nicht-konvexe Technologiemengen



Das vierte Axiom bezieht sich auf die Möglichkeit, Inputs kostenlos zu verschwenden, d.h. ohne Outputreduktion entsorgen zu können. In der strengen Form verlangt dieses Axiom, dass der zusätzliche Einsatz von einzelnen oder mehreren Inputs den Output nie reduziert: Es ist immer möglich, von einem oder mehreren Inputs mehr zu verwenden, ohne dabei die Produktion der Outputs einzuschränken. Dieses Axiom bedingt lediglich, dass die Inputs einen nichtnegativen Wert haben und folglich kostenlos veräußert werden können, daher nicht in den Produktionsprozess eingehen und folglich diesen nicht behindern. Diese Situation wird in Abbildung 2.6a) anhand der Isoquante durch die Inputvektoren A , D und C illustriert, wobei die senkrechten und waagrechten Abschnitte oberhalb Punkt A bzw. rechts von Punkt C ebenfalls zur Isoquante gehören. Die schwache Form verlangt, dass ein gegebener Outputvektor immer mit einem vielfachen des Inputsets hergestellt werden kann. Wie wenig restriktiv diese Annahme ist, soll mit folgende Beispiel erläutert werden: Ist die Produktion eines Outputvektors Y mit einem bestimmten

Inputvektor X möglich, muss – gemäss Axiom vier in der schwachen Form – Y auch mit den doppelten Mengen des Inputvektors hergestellt werden können. Dies ist grundsätzlich möglich, wenn die überschüssigen Inputs ($2X - X$) kostenlos entsorgt werden können. Da $2X - X$ gleich X ist und X die Herstellung des Outputvektors Y ermöglicht, sollte – sofern Y selbst einen Wert hat – die kostenlose Veräusserung von X immer möglich sein. Die

Abbildung 2.6: Kostenlose Entsorgung von Inputs und Outputs



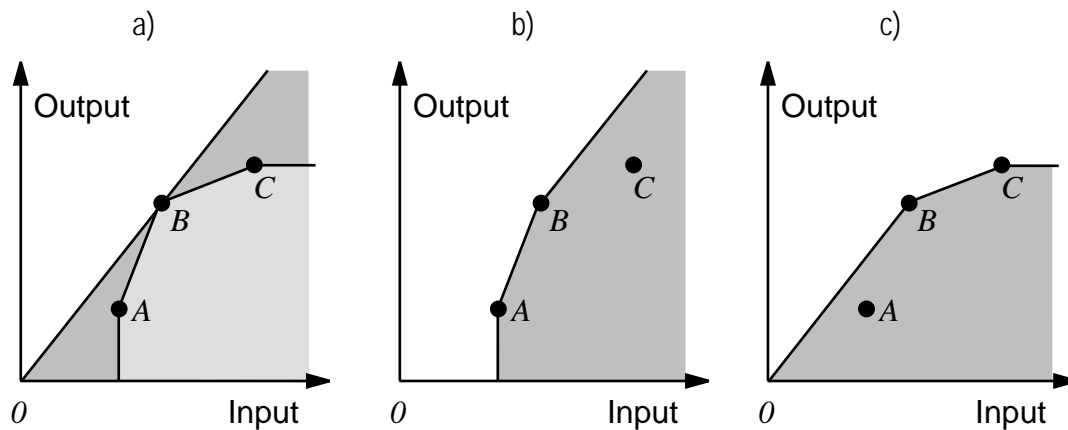
schwache Form dieses Axioms verlangt, dass die Technologiemenge durch die gestrichelten Strahlen in Abbildung 2.6a) durch die Punkte A , B , C und D definiert wird (wobei zu beachten ist, dass die Verlängerungen dieser Strahlen durch den Ursprung gehen).

Das fünfte Axiom entspricht dem vierten, allerdings auf den Output übertragen. In der starken Form können einzelne oder mehrere Outputs reduziert werden, ohne dass die übrigen Outputmengen beeinflusst werden. Wie bei der kostenlosen Entsorgung der Inputs in der starken Form ist diese Bedingung mit einem nichtnegativen Wert aller Outputs in Verbindung zu bringen. Einzelne Outputs können immer kostenlos veräussert werden. Die Isoquante entspricht dann der durchgezogenen Linie durch die Punkte X , Y und W . Die schwache Form besagt, dass mit einem gegebenen Inputvektor immer ein Bruchteil des Outputsets produziert werden kann (vgl. die gestrichelten Linien in Abbildung 2.6b). Das fünfte Axiom in der schwachen Form ist mit einem negativen Wert mindestens eines Outputs in Verbindung zu bringen. Fällt z.B. bei der Produktion von Strom in einem Kohlekraftwerk der unerwünschte Output „Schadstoffemission“ an, kann eine Reduktion der Schadstoffemission um einen bestimmten Prozentsatz erreicht werden,

wenn die Stromproduktion um denselben Prozentsatz reduziert wird.

Das sechste Axiom definiert die Art der Skalenerträge. Grundsätzlich soll mit diesen Annahmen die „wahre“ Technologiemenge so gut wie möglich approximiert werden. Über die Annahmen bezüglich der Skalenerträge muss in der empirischen Forschung von Fall zu Fall entschieden werden.³ In Abbildung 2.7a) ist die Technologiemenge mit konstanten (dunkelgraue Fläche) und mit variablen (hellgraue Fläche) Skalenerträgen dargestellt. In Abbildung 2.7b) ist für dieselben Beobachtungen die Technologiemenge mit nichtabnehmenden Skalenerträgen; in Abbildung 2.7c) ist die Technologiemenge mit nichtzunehmenden Skalenerträgen dargestellt. Je restriktiver die Annahmen sind, desto grösser

Abbildung 2.7: Die Art der Skalenerträge



wird die Technologiemenge sein. Die Technologiemenge mit variablen Skalenerträgen ist eine Teilmenge der Technologiemengen mit nichtzunehmenden bzw. nichtabnehmenden Skalenerträgen. Diese wiederum sind Teilmengen der Technologiemenge mit konstanten Skalenerträgen:

$$\mathcal{T}^{VRS} \subseteq \mathcal{T}^{NIR, NDR} \subseteq \mathcal{T}^{CRS}.$$

³Bewährt hat sich die Vorgehensweise, die Strukturannahmen bezüglich der Skalenerträge schrittweise zu lockern. Einerseits steht mehr Information zur weiteren Auswertung bereit und andererseits erlaubt diese Vorgehensweise, zu untersuchen, unter welchen Umständen sich unterschiedliche Annahmen auf die Effizienz auswirken. Ein Beispiel hierfür ist die Aufteilung der Ineffizienz in rein technische Ineffizienz und in Skaleneffizienz.

Das siebte Axiom verlangt, dass die Technologiemenge empirisch vollständig ist. Alle beobachteten Input-Outputtransformationen gehören zur Technologiemenge: Wenn eine Input-Outputtransformation realisiert wird, muss diese auch zur Technologiemenge gehören. Dieses Axiom kann wie folgt formuliert werden: $Prob[(X_j, Y_j) \in \mathcal{T}] = 1$, $\forall j = 1 \dots n$. Diese trivial erscheinende Aussage ist bei den parametrischen Methoden, die statistische Ausreisser zulassen, nicht erfüllt. Dieses Axiom schliesst ein hier nicht explizit erwähntes Axiom mit ein, welches verlangt, dass ohne Inputs nie Outputs produziert werden können (*no free lunch*).

Das achte Axiom – die minimale Extrapolation der Technologiemenge – verlangt, dass die durch die Technologiemenge definierte Teilmenge die kleinstmögliche Teilmenge im Input-Outputraum sein soll [vgl. Banker, Charnes und Cooper (1984)]. Die effiziente Grenze soll die Daten so eng wie möglich umhüllen.

Die Liste der Axiome ist nicht vollständig; es ist durchaus denkbar, weitere Axiome zu definieren. Es ist auch nicht zwingend, dass die Axiome alle erfüllt sein müssen, denn durch die Wahl der Axiome werden die Eigenschaften der Technologiemenge definiert, die je nach Fragestellung unterschiedlich plausibel, wichtig und wünschenswert sein können.

Zu erwähnen ist an dieser Stelle, dass beim parametrischen Ansatz durch eine explizite, vollständige⁴ oder partielle⁵ Festlegung der Transformationsfunktion in Gleichung 2.6 die geometrische Struktur der Technologiemenge determiniert wird. Durch die Wahl der Transformationsfunktion wird implizit entschieden, welche Axiome erfüllt sind, bzw. die Entscheidung für eine Funktion richtet sich nach den zu erfüllenden Axiomen. Ein Beispiel soll dies illustrieren. Folgende *Cobb-Douglas*-Transformationsfunktion repräsentiere die Technologiemenge:

$$y - x_1^\alpha x_2^\beta \leq 0, \alpha, \beta > 0. \quad (2.7)$$

Wie einfach gezeigt werden kann, ist das erste Axiom erfüllt. Das zweite ist – da nur ein Output produziert wird – nicht definiert. Das dritte Axiom ist nur dann erfüllt, wenn die Summe der Parameter α und β kleiner gleich eins ist. Schliesslich genügt die *Cobb-Douglas*-Transformationsfunktion auch den Axiomen vier und fünf in der starken Form. Das sechste Axiom ist – für konstante Skalenerträge – nur dann erfüllt, wenn die Summe der Parameter α und β gleich eins ist.

⁴Die Funktion inklusive aller Parameter sind festgelegt; die Form der Technologiemenge ist durch die Wahl der Funktion vollumfänglich definiert.

⁵Der funktionelle Zusammenhang ist festgelegt, einige Parameter der Funktion werden geschätzt.

Folgerung 2.1 *Im Rahmen der Effizienzmessung stellen die Firmen Black Boxes dar, deren Leistung darin besteht, in einem bestimmten Zeitraum Inputs in Outputs zu transformieren. Die Produktionsmöglichkeiten werden durch die Technologiemenge repräsentiert, welche empirisch ermittelt oder theoretisch festgelegt werden kann. Die Eigenschaften der Technologiemenge müssen durch Annahmen, die die geometrische Struktur der Grenze der Produktionsmöglichkeiten (GPM) definieren, festgelegt werden.*

2.2 Nichtparametrische Methoden der Effizienzmessung

2.2.1 Effizienz, Effizienzmasse und Effizienzwerte

Koopmans (1951) hat aus dem Wohlfahrtsprinzip, das von Pareto (1897) formuliert wurde und besagt, dass eine Situation dann optimal ist, wenn es unmöglich ist, ein Individuum besserzustellen, ohne ein anderes schlechterzustellen, den folgenden Effizienzbegriff abgeleitet: Eine Input-Outputtransformation ist dann pareto-koopmans-effizient bezüglich einer Technologiemenge \mathcal{T} , wenn es unmöglich ist, sich in einem Kriterium zu verbessern (weniger Inputs oder mehr Outputs) ohne sich in einem anderen zu verschlechtern (mehr Inputs oder weniger Outputs).

Zusammen mit der oben definierten Technologiemenge lassen sich nun Effizienzmasse herleiten. Ein Effizienzmass ist eine Funktion, die jeder Input-Outputtransformation unter Berücksichtigung der Technologiemenge \mathcal{T} einen Effizienzwert zuordnet. Folgende Eigenschaften charakterisieren Effizienzmasse:

- Die *Monotonie*: Ein Effizienzmass ist strikt monoton, wenn eine Verbesserung in mindestens einem Kriterium zu einer Erhöhung des Effizienzwertes führt. Das Konzept der Pareto-Koopmans-Effizienz bedingt zwingend strikt monotone Effizienzmasse.

Ein Effizienzmass wird als schwach monoton in Inputs (Outputs) bezeichnet, wenn eine Reduktion (Erhöhung) *aller* Inputs (Outputs) in einem höheren Effizienzwert resultiert, eine Verbesserung in einzelnen Inputs (Outputs) aber den Effizienzwert unverändert lässt.⁶

⁶Die Monotonie kann auch anhand der Dominanz von Input-Outputtransformationen erläutert werden: Die Dominanz von Input-Outputtransformationen resultiert bei strikt monotonen Effizienzmassen in

- Die *Effizienzindikation*: Ein Effizienzmass ist dann ein Effizienzindikator, wenn anhand des Effizienzwertes pareto-koopmans-effiziente Input-Outputtransformationen erkannt werden können. Implizit verlangt die Effizienzindikation, dass ein Schwellenwert existiert, der nur im Falle pareto-koopmans-effizienter Input-Outputtransformationen erreicht wird. Die Effizienzindikation kann zudem nur bei strikt monotonen Effizienzmassen erfüllt sein.
- Die *Invarianz* bezüglich Datentransformation: Ist ein Effizienzwert unabhängig von der Masseinheit der Inputs und Outputs, so ist das Effizienzmass invariant bezüglich der Transformation der Inputs und Outputs. Von Interesse ist dabei vor allem die Invarianz bezüglich der linearen Transformation. Sie garantiert, dass der Effizienzwert unabhängig von der Masseinheit – z.B. Kilogramm oder Tonne – der Inputs und Outputs ist.

Diese Liste der Eigenschaften von Effizienzmassen ist nicht abgeschlossen. Je nach Fragestellung können weitere Eigenschaften wünschenswert erscheinen. Cooper, Park und Pastor (1999) formulieren im Rahmen des *Range Adjusted Measures* weitere Anforderungen an ein „ideales“ Effizienzmass, wobei bislang kein Effizienzmass entwickelt worden ist, das alle vorgeschlagenen Eigenschaften vereint.⁷ Die Monotonie, die Effizienzindikation und die Invarianz sind aber die in diesem Kontext relevanten Eigenschaften und dienen im Folgenden als Beurteilungskriterien für die hergeleiteten Effizienzmasse.

2.2.2 Die Distanzfunktion

Die oben aufgeführte Definition der Pareto-Koopmans-Effizienz lässt eine Unterteilung der Grenze der Produktionsmöglichkeiten in effiziente, schwach-effiziente und ineffiziente Bereiche zu. Die Inputisoquanten sind definiert durch

$$\text{Iso}\mathcal{L}(Y) = \{X \mid X \in \mathcal{L}(Y) \mid \theta X \notin \mathcal{L}(Y) \forall 0 < \theta < 1\} \quad (2.8)$$

einem höheren Effizienzwert. Bei schwacher Monotonie der Effizienzmasse führt dagegen nur die strikte Dominanz zu einem höheren Effizienzwert.

⁷Drei weitere Eigenschaft werden gefordert: 1) Der Wertebereich der Effizienz ist auf ein Intervall zwischen null bis eins beschränkt, 2) die Input-Outputtransformationen können konsistent nach den Effizienzwerten geordnet werden; 3) das Effizienzmass ist Translationsinvariant, d.h. die Addition von Konstanten zu einzelnen Inputs und Outputs hat keine Konsequenzen auf den Effizienzwert. Dass nicht alle Eigenschaften beim *Range Adjusted Measure* simultan erfüllt sind, zeigen Steinmann und Zweifel (2001).

Eine *proportionale* Reduktion *aller* Inputs X mit dem Faktor θ hat unweigerlich zur Folge, dass die reduzierten Inputs nicht mehr im Inputset $\mathcal{L}(Y)$ enthalten sind (vgl. Gleichung 2.8). Die Inputisoquante stellt die GPM dar, garantiert aber noch nicht Pareto-Koopmans-Effizienz. Die effizienten Abschnitte sind diejenigen Punkte der Inputisoquanten, bei denen eine Reduktion *einzelner* Inputs zu einer Reduktion der Outputs führt und ist definiert durch⁸

$$\text{Eff}\mathcal{L}(Y) = \{X \mid X \in \mathcal{L}(Y) \mid \bar{X} \notin \mathcal{L}(Y) \forall \bar{X} \leq X\}. \quad (2.9)$$

Die Isoquante setzt sich demzufolge aus pareto-koopmans-effizienten Abschnitten, gemäss Gleichung 2.9, und schwach effizienten Abschnitten, die parallel zu den Achsen verlaufen, zusammen.

Die Outputisoquante ist dementsprechend durch

$$\text{Iso}\mathcal{P}(X) = \{Y \mid Y \in \mathcal{P}(X) \mid \theta^{-1}Y \notin \mathcal{P}(X) \forall 0 < \theta < 1\} \quad (2.10)$$

definiert. Eine *proportionale* Ausdehnung *aller* Outputs Y mit dem Faktor θ^{-1} ist nicht möglich und hat zur Folge, dass Y nicht mehr im Outputset $\mathcal{P}(X)$ enthalten ist. Die effizienten Abschnitte der Outputisoquanten sind diejenigen Punkte, die keine Ausdehnung *einzelner* Outputs bei gegebenem Inputvektor zulassen

$$\text{Eff}\mathcal{P}(X) = \{Y \mid Y \in \mathcal{P}(X) \mid \bar{Y} \notin \mathcal{P}(X) \forall Y \leq \bar{Y}\}. \quad (2.11)$$

Für jeden beliebigen Inputvektor $X \in \mathcal{L}(Y)$ und jeden beliebigen Outputvektor $Y \in \mathcal{P}(X)$ lässt sich gemäss Farrell (1957) ein Mass der *technischen Effizienz* anhand der folgenden Distanzfunktionen bestimmen:

$$\theta(Y_i, X_i) = \inf\{\theta \mid \theta X_i \in \mathcal{L}(Y_i)\} = \inf\{\theta \mid (Y_i, \theta X_i) \in \mathcal{T}\} \quad (2.12)$$

$$\theta(Y_i, X_i) = \inf\{\theta \mid \theta^{-1} Y_i \in \mathcal{P}(X_i)\} = \inf\{\theta \mid (\theta^{-1} Y_i, X_i) \in \mathcal{T}\}. \quad (2.13)$$

Ist θ gleich eins, liegt der Input- bzw. der Outputvektor genau auf der Isoquante, ist θ dagegen kleiner als eins, lässt sich der Inputvektor [Outputvektor] *proportional* reduzieren

⁸Zur Vermeidung von Missverständnissen werden die Ungleichheitszeichen im Zusammenhang mit Vektoren definiert: 1) $a > b$ bedeutet $a_n > b_n$ für alle n Elemente von a und b ; 2) $a \geq b$ bedeutet $a_n \geq b_n$ für alle n und $a_n > b_n$ für einige n Elemente von a und b und 3) $a \geq b$ bedeutet $a_n \geq b_n$ für alle n Elemente. Ist $b = 0$, wird a in 1) als positiv, in 2) als semipositiv und in 3) als nichtnegativ bezeichnet.

[ausdehnen] bis er auf der Isoquante zu liegen kommt. Jede weitere Reduktion [Ausdehnung] hat zur Folge, dass X [Y] nicht mehr in $\mathcal{L}(Y)$ [$\mathcal{P}(X)$] enthalten ist. Zu beachten bleibt, dass die proportionale Änderung der Inputs und Outputs nicht zwingend eine Projektion auf die effiziente Grenze zur Folge hat; es ist auch möglich, dass ineffiziente Input-Outputtransformationen auf die schwach effiziente Grenze projiziert werden.⁹ Das Distanzmass kann als Effizienzmass interpretiert werden: Je kleiner θ ist, desto mehr Inputs lassen sich bei gegebenem Output einsparen bzw. desto mehr Outputs lassen sich bei gegebenen Inputs produzieren.¹⁰

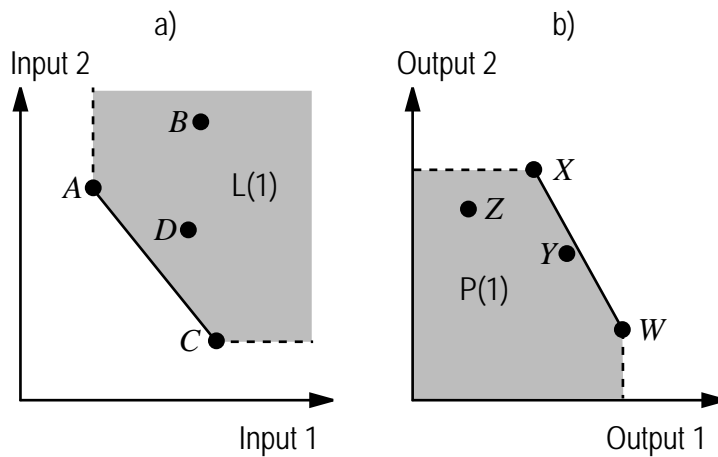
In Abbildung 2.8 ist die Technologiemenge dargestellt, die den Axiomen 1 bis 8 genügt. Die Outputisoquante des Inputsets $\mathcal{L}(Y)$ für den Fall mit zwei Inputs ist in Abbildung 2.8a) dargestellt. Wie viele verschiedene Güter Y umfasst ist in diesem Beispiel irrelevant, Y könnte ein einzelnes oder auch ein Bündel von verschiedenen Gütern sein. Die Strecke AC der Isoquante ist effizient, eine Reduktion eines einzelnen Inputs ist bei gegebenem Outputvektor nicht möglich. Die gestrichelten Abschnitte der Isoquante, die parallel zu den Achsen verlaufen, sind schwach-effizient. Eine proportionale Reduktion aller Inputs ist bei den zu produzierenden Outputmengen nicht zulässig. In Abbildung 2.8b) ist die Inputisoquante des Outputsets $\mathcal{P}(X)$ dargestellt, die beiden Outputs werden dabei mit einem gegebenen Inputvektor hergestellt. Die Strecke XW ist, gemäss Gleichung 2.11, der effiziente Abschnitt der Isoquante.

Abschliessend bleibt zu erwähnen, dass die effiziente Grenze bei der nichtparametrischen Technologiemenge im Gegensatz zur parametrischen Technologiemenge ausschliesslich durch *Best-Practice*-Input-Outputtransformationen, d.h. durch effiziente Beobachtungen, definiert wird. Dies ist ein Vorteil gegenüber den parametrischen Technologiemengen, da die ineffizienten Input-Outputtransformationen die GPM nicht beeinflussen und folglich nicht verzerren können. Andererseits ist zu beachten, dass die durch *Best Prac-*

⁹Bei einer Beobachtung, die auf die schwach effiziente Grenze projiziert wird, können offensichtlich einzelne Inputs ohne Outputrückgang reduziert werden. Diese überschüssigen Inputs werden Inputslacks genannt. Analog sind die Outputslacks definiert: Sie liegen vor, wenn von einigen Outputs mehr produziert werden können, ohne dabei zusätzliche Inputs zu benötigen.

¹⁰Farrell's Mass der technischen Effizienz fand lange Zeit keinen grossen Anklang in der empirischen Forschung. Der Grund dafür ist, dass kein effizienter Lösungsalgorithmus bekannt war. Zur Berechnung der Effizienz wurden für jede Beobachtung *alle* möglichen Kombinationen von Beobachtungen evaluiert und dann diejenige bestimmt, bei der θ den kleinsten Wert annahm. Diese Prozedur erfordert die Lösung von $N = (n + m + s)!/n!(m + s)!$ linearen Gleichungssystemen (d.h. ebenso viele Matrixinversionen und anschliessende -multiplikationen).

Abbildung 2.8: Die Technologiemenge



tices definierte Technologie sensitiv auf Messfehler und stochastische Einflüsse reagiert. Solche Fehler in den Daten haben zur Folge, dass die Effizienz *ceteris paribus* unterschätzt wird, da die GPM nur nach „aussen“, d.h. in Richtung weniger Inputs und mehr Outputs, verschoben werden kann.

2.2.3 Die Data Envelopment Analysis (DEA)

Die DEA wurde von Charnes et al. (1978) eingeführt und stellt eine Weiterentwicklung der von Farrell (1957) vorgeschlagenen Methode zur Berechnung der *technischen Effizienz* dar. Die DEA kann in diesem Sinne auch als Distanzfunktion bezeichnet werden. Die Arbeit von Charnes et al. (1978) war insofern innovativ, als dass sie einen Algorithmus in der Form eines Linearen Programms zur Lösung des von Farrell gestellten Problems formulierten.

In diesem Abschnitt wird die inputorientierte DEA-Version hergeleitet und ausführlich dargestellt. Inputorientiert bedeutet in diesem Zusammenhang, dass gemäß Gleichung 2.12 die maximal zulässige radiale Kontraktion des Inputvektors bei gegebenem Outputvektor gesucht wird. Weiterhin gelten die in Abschnitt 2.1 aufgeführten Axiome 1 bis 8. Die Technologie ist konvex (Axiome 1 bis 3) und hat die Eigenschaft der freien Entsorgung der Inputs und Outputs in der starken Form (Axiome 4 und 5). Weiter werden konstante Skalenerträge unterstellt (Axiom 6a) und die Technologiemenge ist empirisch

vollständig (Axiom 7). Folglich ist diese Technologiemenge die kleinstmögliche Teilmenge im Input-Outputraum (Axiom 8), die den Axiomen 1 bis 7 genügt.

Das ursprüngliche Lineare Programm: Das Multiplier-Problem

Die zugrunde liegende Idee ist die Verallgemeinerung der eindimensionalen Produktivität.¹¹ Wird die Produktivität verschiedener Beobachtungen miteinander verglichen, ist ein grosser Quotient mit hoher Effizienz gleichzusetzen. Der Absolutwert der Produktivität hat aber isoliert betrachtet keine Aussagekraft. Erst in Relation zur Produktivität anderer Beobachtungen können Aussagen zur *relativen* Effizienz gemacht werden.

Das Ziel der Verallgemeinerung der Produktivitätskennzahl besteht darin, den Quotienten der *gewichteten* Outputs und Inputs unter Nebenbedingungen zu maximieren. Ausgangslage ist das folgende Nichtlineare Programm in Gleichung 2.14. Y_l ist der Outputvektor mit der Dimension $s \times 1$ und X_l der Inputvektor mit der Dimension $m \times 1$ der zu evaluierenden Beobachtung l .¹² Die durch das Nichtlineare Programm zu bestimmenden Gewichte der Outputs werden in einem Vektor ν , diejenigen der Inputs in einem Vektor μ zusammengefasst. Die Elemente der Vektoren ν und μ werden in der DEA-Terminologie *virtuelle Multiplikatoren* genannt und können als Schattenpreise interpretiert werden. Sie stellen die Grundlage für die Aggregation der Inputs und Outputs dar.¹³

$$\begin{aligned} \max \quad & h = \frac{\nu^T Y_l}{\mu^T X_l} \\ \text{s.t.} \quad & \frac{\nu^T Y_i}{\mu^T X_i} \leq 1 \quad \forall i = 1, \dots, l, \dots, n \\ & \nu, \mu \geq 0 \end{aligned} \tag{2.14}$$

Die n Nebenbedingungen – eine für jede Beobachtung – verlangen, dass kein Quotient einen Wert grösser eins annimmt und dienen der Normierung; h kann folglich maximal den Wert von eins annehmen.¹⁴ Eine Beobachtung l ist dann effizient, wenn virtuelle Multiplikatoren existieren, so dass keine andere Beobachtung mit eben diesen virtuellen

¹¹Die Produktivität ist definiert als der Quotient zwischen einem Output und einem Input.

¹²Allgemein wird angenommen, dass die Daten den folgenden Anforderungen genügen: 1) Jeder Input wird mindestens von einer Firma verwendet; 2) jeder Output wird mindestens von einer Firma produziert; 3) jede Firma verwendet mindestens einen Input und produziert mindestens einen Output und 4) identische Input-Outputtransformationen werden nur einmal in die Analyse aufgenommen.

¹³Das Superskript T deutet darauf hin, dass es sich um einen transponierten Vektor handelt.

¹⁴Die zu evaluierende Beobachtung j stellt ja die j^{te} Nebenbedingung.

Multiplikatoren einen grösseren Quotienten aufweist und folglich die Nebenbedingung l gleich eins ist.

Dieses Verfahren zur Bestimmung der Gewichte der Outputs und Inputs ist insofern als *fair* zu bezeichnen, als dass keine von der optimalen Lösung verschiedene virtuelle Multiplikatoren – oder Gewichte bzw. Preise – existieren, die zu einer höheren Effizienz führen. Jede Beobachtung wird sozusagen ins „beste Licht“ gerückt.

Die Zielfunktion in Gleichung 2.14 ist nichtlinear und es existieren unendlich viele Lösungen: Stellen μ^* und ν^* eine Lösung zum Problem in Gleichung 2.14 dar, so ist auch jedes Vielfache von μ^* und ν^* wieder eine Lösung. Das Nichtlineare Programm kann aber in ein Lineares Programm transformiert werden, wenn der Nenner in der Zielfunktion auf eins normiert wird, indem eine zusätzliche Nebenbedingung

$$\mu^T X_l = 1$$

eingeführt wird. Dies hat zur Folge, dass der Nenner eliminiert werden kann. Damit auch die n Nebenbedingungen linear sind, müssen diese mit $\mu^T X_i \forall i = 1, \dots, l, \dots, n$ multipliziert werden.¹⁵ Das so transformierte Lineare Programm ist in Gleichung 2.15 aufgeführt. Die n Outputvektoren werden, damit die übersichtlichere Matrixschreibweise verwendet werden kann, zur Outputmatrix \mathbf{Y} der Dimension $s \times n$ und die n Inputvektoren zur Inputmatrix \mathbf{X} der Dimension $m \times n$ zusammengefasst.¹⁶ Die Zielfunktion nimmt denselben Wert an wie das Nichtlineare Programm ($h^* = z^*$), besitzt aber im Gegensatz zu diesem eine eindeutige Lösung für die virtuellen Multiplikatoren ν^* und μ^* .¹⁷

$$\begin{aligned} \max \quad & z = \nu^T Y_l \\ \text{s.t.} \quad & \mu^T X_l = 1 \\ & \mathbf{Y}^T \nu - \mathbf{X}^T \mu \leq 0 \\ & \nu, \mu \geq 0 \end{aligned} \tag{2.15}$$

Das Lineare Programm in Gleichung 2.15 muss für jede Beobachtung, d.h. insgesamt n -mal, gelöst werden,¹⁸ wobei die virtuellen Multiplikatoren die Variablen¹⁹ und die beiden

¹⁵Dies ist immer zulässig, da alle Elemente in X und μ auf nichtnegative Werte beschränkt sind und folglich $\mu^T X_i$ auch nichtnegativ sein muss.

¹⁶Werden die Input- und die Outputmatrix zusammengefasst spricht man von der Input-Outputmatrix. Diese ist in diesem Kontext $(m + s) \times n$ -dimensional.

¹⁷Ein Stern deutet darauf hin, dass es sich um eine Lösung zum Linearen Programm handelt.

¹⁸Lineare Programme können mit dem *Simplex*-Algorithmus gelöst werden, der auf Danzig (1951) zurückgeht. Eine übersichtliche Einführung findet sich in Chiang (1984).

¹⁹Daher wird dieses Lineare Programm auch als *Multiplier-Problem* bezeichnet.

Vektoren μ^* und ν^* die Lösung des Maximierungsproblems sind.

Neben den Nichtnegativitätsbedingungen der virtuellen Multiplikatoren müssen wieder n Nebenbedingungen erfüllt sein. Die Summe der gewichteten Outputs jeder Beobachtung minus die Summe der gewichteten Inputs muss kleiner gleich null sein. Zur Illustration sind diese Nebenbedingungen für $n = 4$ Beobachtungen, mit jeweils $m = 2$ Inputs und $s = 2$ Outputs explizit dargestellt:

$$\begin{aligned}\nu_1 y_{11} + \nu_2 y_{12} - \mu_1 x_{11} - \mu_2 x_{12} &\leq 0 \\ \nu_1 y_{21} + \nu_2 y_{22} - \mu_1 x_{21} - \mu_2 x_{22} &\leq 0 \\ \nu_1 y_{31} + \nu_2 y_{32} - \mu_1 x_{31} - \mu_2 x_{32} &\leq 0 \\ \nu_1 y_{41} + \nu_2 y_{42} - \mu_1 x_{41} - \mu_2 x_{42} &\leq 0\end{aligned}$$

Jede der $n = 4$ Zeilen kann als Transformationsfunktion der Technologiemenge betrachtet werden. Ist die n^{te} Nebenbedingung bei gegebenen virtuellen Multiplikatoren bindend, d.h. gleich null, ist die entsprechende Beobachtung effizient. Wird z.B. die erste Beobachtung evaluiert und sind die dritte und vierte Gleichung mit den das Problem in Gleichung 2.15 maximierenden virtuellen Multiplikatoren bindend, definieren die dritte und vierte Beobachtung den Abschnitt der GPM, an der die erste Beobachtung evaluiert wird.²⁰

Aus den Nebenbedingungen lässt sich unmittelbar der Wertebereich der Zielfunktion bestimmen. Da die Outputmengen und die virtuellen Multiplikatoren auf nichtnegative Werte beschränkt sind, kann die Summe der gewichteten Outputs keine negativen Werte annehmen. Bei absolut ineffizienten Beobachtungen – wenn keine Outputs produziert werden – wird die Zielfunktion demzufolge einen Wert von null annehmen. Die Zielfunktion einer effizienten Beobachtung i wird hingegen den Wert eins annehmen, da die i^{te} Nebenbedingung bindend und die Summe der gewichteten Inputs gleich eins sein muss. Der Wert z^* der Zielfunktion des Linearen Programms muss folglich für eine ineffiziente Beobachtung j einen Wert zwischen null und eins annehmen. Die j^{te} Nebenbedingung ist dann nicht-bindend

$$\underbrace{(\nu_1 y_{j1} + \nu_2 y_{j2})}_{z^*} - \underbrace{(\mu_1 x_{j1} + \mu_2 x_{j2})}_1 = < 0.$$

Zur Illustration der Bedeutung der optimalen Lösung z_j^* wird nun der ineffiziente Inputvektor X_j mit dem Skalar z_j^* multipliziert und das Lineare Programm für die j^{te}

²⁰Die diesen Abschnitt der GPM definierenden Beobachtungen werden als Referenzset für die evaluierte Beobachtung bezeichnet.

Beobachtung wiederholt. Damit die Summe der gewichteten Inputs mit dem reduzierten Inputvektor wieder gleich eins ist (vgl. die erste Nebenbedingung in Gleichung 2.15), muss der ursprüngliche Vektor der virtuellen Multiplikatoren μ^* mit dem reziproken Wert von z_j^* multipliziert werden. Da die Input- und Outputmengen der dritten und vierten Beobachtung unverändert bleiben und die entsprechenden Nebenbedingungen bindend bleiben müssen, muss auch der Vektor der virtuellen Multiplikatoren der Outputs ν^* mit demselben Wert (z^{*-1}) multipliziert werden. Die j^{te} Nebenbedingung lautet dann

$$z^* \cdot z^{*-1} - 1 = 0. \quad (2.16)$$

Die j^{te} Nebenbedingung ist nun, wie aus Gleichung 2.16 ersichtlich ist, bindend. Dies bedeutet nichts anderes, als dass die neue Input-Outputtransformation auf der GPM liegt, bzw. auf diese projiziert wurde. Die optimale Lösung z^* in Gleichung 2.15 entspricht der maximal zulässigen, radialen Reduktion des Inputvektors und ist gleich θ^* aus Gleichung 2.12.

Die beiden Vektoren der virtuellen Multiplikatoren einer Lösung des Linearen Programms in Gleichung 2.15 definieren eine Ebene im $m + s$ -dimensionalen Raum (ohne Einschränkung der Allgemeinheit wird hier der 4-dimensionale Fall dargestellt):

$$\nu_1 y_1 + \nu_2 y_2 - \mu_1 x_1 - \mu_2 x_2 = 0. \quad (2.17)$$

Wird Gleichung 2.17 total differenziert, lassen sich alle Grenzzraten der Substitution und Transformation sowie die Grenzproduktivitäten bestimmen. Somit verfügen alle Beobachtungen, die auf dieser Ebene liegen oder im Falle der ineffizienten Beobachtungen auf diese Ebene projiziert werden, über die gleiche Transformationsfunktion:

$$\begin{aligned} &\nu_1 dy_1 + y_1 d\nu_1 + \nu_2 dy_2 + y_2 d\nu_2 \\ &- \mu_1 dx_1 - x_1 d\mu_1 - \mu_2 dx_2 - x_2 d\mu_2 = 0. \end{aligned} \quad (2.18)$$

Das duale Lineare Programm: Das Envelopment-Problem

Das duale Lineare Programm²¹ zum *Multiplier-Problem* in Gleichung 2.15 ist in Gleichung 2.19 aufgeführt und wird als *Envelopment-Problem* bezeichnet. Das θ , das es zu minimieren gilt, und die n Elemente des Vektors λ sind die Variablen; die Input-Outputmatrix

²¹Für jedes Lineare Programm existiert eine duales Lineares Programm. Aus einem Maximierungsproblem wird ein Minimierungsproblem, die Koeffizientenmatrix (die Input-Outputmatrix) wird transponiert und die Elemente der Zielfunktion werden zu den Variablen auf der rechten Seite der Nebenbedingungen und *vice versa*. Zudem werden die Ungleichheitszeichen gedreht [vgl. z.B. Chiang (1984), Kapitel 20].

ist die Koeffizientenmatrix. Da n in der Regel viel grösser als $m + s$ ist, kommt das *Envelopment-Problem* mit $n - (m + s)$ weniger Nebenbedingungen aus und ist folglich mit geringerem Aufwand zu lösen.

$$\begin{aligned}
 & \min \quad \theta \\
 \text{s.t.} \quad & \theta X_l - \mathbf{X}\lambda \geq 0 \\
 & Y_l - \mathbf{Y}\lambda \leq 0 \\
 & \lambda \geq 0
 \end{aligned} \tag{2.19}$$

Das Ziel des *Envelopment-Problems* ist, eine Linearkombination von Input-Outputtransformationen zu finden, die eine grösstmögliche radiale Kontraktion des zu evaluierenden Inputvektors zulässt; dies ist exakt dieselbe Fragestellung wie in Farrell's Distanzfunktion (vgl. Gleichung 2.12). Die erste Zeile der Nebenbedingungen in Gleichung 2.19 fasst die m Bedingungen zusammen, die verlangen, dass die um den Faktor θ^* reduzierten Inputs nicht kleiner als die Linearkombination der Inputs der Referenzbeobachtungen sind. Die zweite Zeile (die s Ungleichungen umfasst) stellt sicher, dass auch wirklich der Outputvektor der zu evaluierenden Beobachtung hergestellt wird. Diese $m + s$ Nebenbedingungen garantieren, dass der um den Faktor θ^* reduzierte Inputvektor und der Outputvektor innerhalb der Technologiemenge liegen.

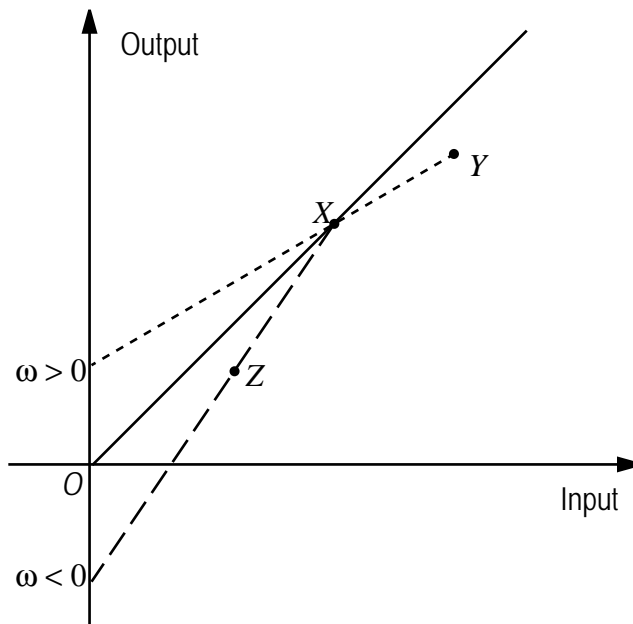
Bei der Lösung des Linearen Programms in Gleichung 2.19 werden die $m + s$ Nebenbedingungen in Gleichungen transformiert, indem jeder Ungleichung eine (nicht-negative) Schlupfvariable zugeordnet wird. Nimmt die Schlupfvariable eines Inputs e im Optimum, d.h. bei minimalem θ^* , einen Wert grösser null an, besteht sogenannter *Inputslack*. Eine zusätzliche proportionale Reduktion aller Inputs ist nicht möglich, doch in Bezug auf den Input e lässt sich eine Verbesserung erzielen, ohne dass der Output reduziert werden müsste. Positive *Inputslacks* oder *Outputslacks* (wie die Schlupfvariablen für Outputs genannt werden) zeigen an, dass eine Beobachtung schwach effizient ist bzw. auf einen schwach effizienten Abschnitt der GPM projiziert wird (im Vergleich zur Pareto-Koopmans-Effizienz, vgl. Abschnitt 2.2.2).

Das erste Theorem der Dualität verlangt, dass der Wertebereich der Variable θ gleich dem Wertebereich der dualen Zielfunktion z sein muss [vgl. z.B. Chiang (1984), S. 690 ff.]. Die optimale Lösung in Gleichung 2.19 für eine effiziente Beobachtung muss folglich den Wert eins annehmen. Weist eine Beobachtung keine von null verschiedenen Mengen an Outputs auf, wird sie mit dem Ursprung des Input-Outputraumes verglichen und θ^* nimmt den Wert null an.

DEA-Versionen mit variablen Skalenerträgen

Zuweilen ist die Annahme konstanter Skalenerträge (Axiom 6a) nicht tragbar bzw. unerwünscht. Banker et al. (1984) schlagen deshalb eine Anpassung der Linearen Programme 2.15 und 2.19 vor, die Technologiemengen mit variablen Skalenerträgen ermöglichen. Im *Multiplier-Problem* wird eine (in keiner Weise restringierte) Variable ω in der Zielfunktion und jeder der n Nebenbedingungen additiv eingeführt. Somit besteht die Möglichkeit, dass die effizienten Ebenen nicht durch den Ursprung des Input-Outputraumes verlaufen müssen. Die Variable ω ist als Achsenabschnitt zu interpretieren. Positive Werte von ω^* deuten darauf hin, dass die evaluierte Beobachtung Y in Abbildung 2.9 im Bereich der abnehmenden Skalenerträge operiert. Werte gleich null werden bei Beobachtungen, die im Bereich der optimalen Betriebsgrösse agieren (wie z.B. Beobachtung X) und Werte kleiner null bei Beobachtungen, die im Bereich der zunehmenden Skalenerträge produzieren (vgl. Beobachtung Z), realisiert. Eine zusätzliche Variable in einem Linearen Programm führt

Abbildung 2.9: DEA mit variablen Skalenerträgen



zu einer zusätzlichen Restriktion im dualen Linearen Programm. Variable Skalenerträge werden im *Envelopment-Problem* mit der zusätzlichen Nebenbedingung $\sum_{j=1}^n \lambda_j^* = 1$ erreicht. Die intuitive Interpretation dieser Restriktion ist, dass der Referenzpunkt der zu

evaluierenden Beobachtung eine konvexe Linearkombination von effizienten Beobachtungen sein muss. Der Ursprung des Input-Outputraumes wird nur noch in Ausnahmefällen ein Element der effizienten Abschnitte sein.²² In Abbildung 2.9 ist die unter der Annahme konstanter Skalenerträge ineffiziente Beobachtung Y effizient, es existiert keine konvexe Linearkombination anderer Beobachtungen, die eine Reduktion der Inputs zulässt. Dasselbe gilt für die Beobachtung Z .

DEA-Versionen mit restringierter Grenze der Produktionsmöglichkeiten

Wie oben erwähnt, können die virtuellen Multiplikatoren aus ökonomischer Sicht als Schattenpreise interpretiert werden. Die Verhältnisse der Schattenpreise entsprechen 1) den Grenzkosten der Substitution (wenn Inputs verglichen werden), 2) der marginalen Produktivität (wenn jeweils ein Input und ein Output verglichen wird) und 3) der Grenzrate der Transformation (wenn Outputs verglichen werden). Thompson, Langemeier, Lee, Lee, und Thrall (1990) und Charnes, Cooper, Huang und Sun (1990) haben Erweiterungen der DEA vorgeschlagen, die auf dieser Erkenntnis aufbauen und die Integration von Preisinformationen in die DEA zulassen. Wären alle Preise bekannt, liesse sich neben der technischen Effizienz und der Skaleneffizienz auch die allokativen Effizienz bestimmen, die dritte Komponente des ökonomischen Effizienzkonzepts. Typischerweise sind aber gerade diese Preisinformationen nicht oder nur unvollständig vorhanden. Die Grössenordnungen sind jedoch oft bekannt: z.B. ist denkbar, dass der Wert einer Einheit des einen Inputs grösser ist als der Wert eines anderen Inputs. Angenommen es sei bekannt, dass das Wertverhältnis zweier Inputfaktoren e und f zwischen a und b liege, gilt

$$a \leq \frac{\mu_e}{\mu_f} \leq b \Rightarrow \begin{cases} -\mu_e + a\mu_f \leq 0 \\ +\mu_e - b\mu_f \leq 0. \end{cases} \quad (2.20)$$

Diese Ungleichungen können auch in Matrixform dargestellt werden und als zusätzliche Nebenbedingungen in das Lineare Programm 2.15 integriert werden

$$\begin{bmatrix} -1 & a \\ 1 & -b \end{bmatrix} \begin{pmatrix} \mu_e \\ \mu_f \end{pmatrix} = \mathbf{C}\mu \leq 0. \quad (2.21)$$

Der Vorteil dieser zusätzlichen Restriktionen besteht darin, dass bei der Lösung der Linearen Programme 2.15 und 2.19 den optimalen virtuellen Multiplikatoren μ^* und ν^* keine

²²Im Modell mit konstanten Skalenerträgen ist der Ursprung dagegen ein Element *aller* Abschnitte der GPM.

aus allokativer Sicht unvernünftigen Werte zugewiesen werden. Nur Beobachtungen, die diesen zusätzlichen Restriktionen genügen, können auch ökonomisch effizient sein. Zudem kann verhindert werden, dass virtuelle Multiplikatoren gleich null realisiert werden. Absorbiert eine Beobachtung Inputs bzw. produziert sie Outputs, muss davon ausgegangen werden, dass diese im Rahmen des Transformationsprozesses bzw. auf den Absatzmärkten einen Wert haben. Intuitiv leuchtet es nicht ein, dass unter dieser Annahme einzelnen Inputs und Outputs bei der Lösung der DEA Schattenpreise gleich null zugewiesen werden.

DEA-Versionen mit restringierten Grenzkosten der Substitution und/oder Transformation werden als *Cone-Ratio-DEA* bezeichnet. Der Begriff *Cone-Ratio* ist auf die Tatsache zurückzuführen, dass diese zusätzlichen Restriktionen jeweils einen konvexen Kegel der zulässigen virtuellen Multiplikatoren im positiven Orthanten des μ - bzw. ν -Raum (Input- bzw. Outputpreisraum) definieren.²³ *Cone-Ratio*-Modelle haben immer eine Lösung.

Werden zudem die Grenzproduktivitäten restringiert, spricht man von der *Linked-Ratio-DEA*, da die zulässigen Bereiche der virtuellen Multiplikatoren der beiden μ - bzw. ν -Räume verbunden werden. *Linked-Ratio*-Modelle haben unter Umständen keine Lösung. Dies ist der Fall, wenn konstante Skalenerträge unterstellt werden und die Restriktionen verhindern, dass mindestens ein effizienter Abschnitt durch den Ursprung verlaufen kann, m.a.W. wenn

$$\frac{\nu_{ke}^*}{\mu_{kg}^*} < \frac{y_{ke}}{x_{kg}} \quad \forall e \in (1, \dots, e, \dots, s), \quad \forall g \in (1, \dots, g, \dots, m), \quad \forall k \in (1, \dots, k, \dots, n).$$

Diese Situation ist in Abbildung 2.10a) dargestellt. Der Bereich der zulässigen Grenzproduktivitäten ist grau schattiert; die effiziente Grenze kann nicht gleichzeitig durch den Punkt X und den Ursprung verlaufen, da die Beobachtung X über ein höheres Output-Inputverhältnis verfügt, als die Restriktion der Grenzproduktivität zulässt.

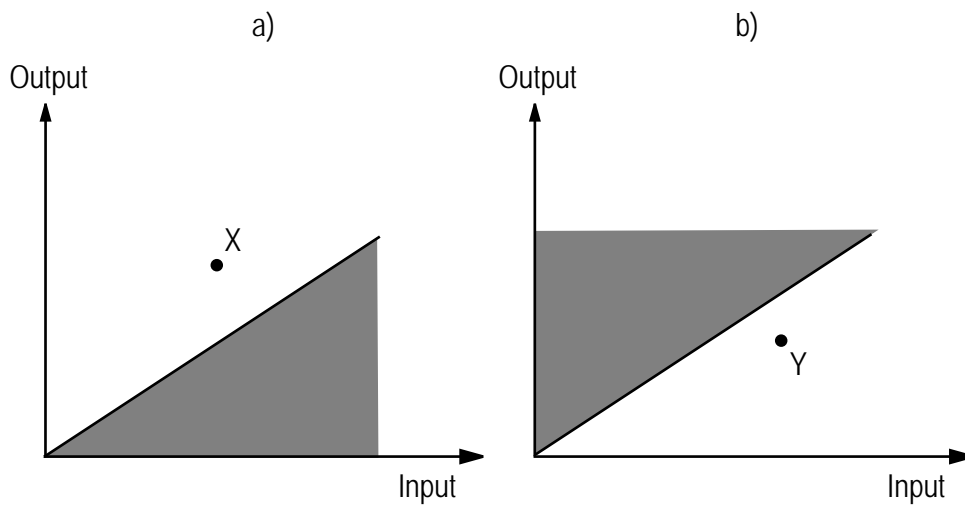
Ist dagegen

$$\frac{\nu_{ke}^*}{\mu_{kg}^*} > \frac{y_{ke}}{x_{kg}} \quad \forall e \in (1, \dots, e, \dots, s), \quad \forall g \in (1, \dots, g, \dots, m), \quad \forall k \in (1, \dots, k, \dots, n),$$

existiert für jede Beobachtung zwar eine Lösung zum Linearen Programm, keine einzige Beobachtung wird aber als schwach effizient, und somit auch nicht als effizient nach Pareto-Koopmans, erkannt. In Abbildung 2.10b) sei Y die effizienteste Beobachtung. Der Bereich

²³ Alternativ wird diese DEA-Version auch *Assurance Region Type I* genannt. Diese Bezeichnung deutet darauf hin, dass mit den Restriktionen sichergestellt wird, dass die optimalen Gewichte (d.h. die virtuellen Multiplikatoren) auf bestimmte Werte beschränkt sind.

der zulässigen Grenzproduktivitäten, wieder grau schattiert, ist auf dermassen hohe Werte restringiert, dass nicht einmal die effizienteste Beobachtung als effizient erkannt werden kann.

Abbildung 2.10: *Linked-Ratio-Modelle*

Diese zusätzlichen Restriktionen führen dazu, dass die Technologiemenge grösser wird, d.h. die Daten nicht gleichermassen eng umschlossen werden wie ohne Restriktionen: $\mathcal{T}^{ES} \subseteq \mathcal{T}_{Cone-Ratio}^{ES} \forall ES = \{CRS, NIR, NDR, VRS\}$. Folglich nimmt der Effizienzwert aller Beobachtungen mit den zusätzlichen Restriktionen und der Anteil der als effizient deklarierten Beobachtungen nie zu.

Zu beachten ist zudem, dass *Cone-Ratio*-Modelle, bei denen alle virtuelle Multiplikatoren auf positive Werte beschränkt sind, strikt monotone Effizienzmasse sind und folglich eine Effizienzindikation zulassen.²⁴

Zum Abschluss dieses Abschnittes soll noch kurz auf die statistische Fundierung der DEA eingegangen werden. Bei der Bestimmung der GPM handelt es sich um ein typi-

²⁴Auf Lineare Programme übertragen besagt die *Complementary-Slackness*-Bedingung, dass, wenn eine Variable (ν, μ) im Optimum einen von null verschiedenen Wert annimmt, im dualen Linearen Programm die Schlupfvariable (=Slack) des entsprechenden Inputs bzw. Outputs gleich null sein muss und *vice versa*. In einem *Cone-Ratio*-Modell, bei dem alle virtuellen Multiplikatoren auf positive Werte beschränkt sind, müssen folglich alle Slacks gleich null sein.

ches statistisches Problem: Anhand einer Stichprobe wird die „wahre“ Technologiemenge geschätzt. Trotzdem wurde die DEA von den traditionellen Ökonometrikern stets skeptisch betrachtet, da ihr anscheinend jede statistische Grundlage fehlte. Eine ganze Reihe von Publikationen liefert aber eine statistische Fundierung der DEA (vgl. Banker (1993), Banker (1996), Grosskopf (1996), Kneip, Park und Simar (1998) und Simar und Wilson (2000), zudem wird im Abschnitt 3.1 auf die statistische Fundierung der DEA eingegangen.). In den folgenden Ausführungen wird daher der Begriff der Effizienzschtätzung anstelle der Effizienzmessung verwendet.

2.3 Erweiterungen der DEA

Den oben beschriebenen Distanzmassen haftet gemeinsam an, dass unterschiedliche Rahmenbedingungen bei der Produktion unberücksichtigt bleiben.²⁵ Die Rahmenbedingungen stellen nicht Inputs im eigentlichen Sinne dar, trotzdem beeinflussen sie den Transformationsprozess in einer bestimmten Art. Farrell (1957) erwähnt bei der Berechnung der Effizienz der U.S. Landwirtschaft

...the apparent differences in efficiency... reflect factors like climate, location and fertility that have not been included in the analysis, as well as „genuine“ differences in efficiency.

Das *Benchmarking*, d.h. der Vergleiche der Effizienzwerte zwischen verschiedenen Firmen, ist daher nicht ohne weiteres möglich. Der Rückschluss von der berechneten relativen Distanz zur GPM auf die Effizienz der Leistungserbringung und die den Transformationsprozess bestimmenden innerbetrieblichen Faktoren – wie z.B. die Produktivität der Arbeit und die Fähigkeiten des Managements – ist daher nicht ohne Weiteres zulässig. Insbesondere werden die Entscheidungsträger der Firmen, die relativ schlecht abschneiden, kritisieren, dass ihre spezifischen Rahmenbedingungen eine effizientere Produktion nicht zulassen. Das relativ schlechte Abschneiden wird dann damit entschuldigt, dass die Leistungserbringung unter den gegebenen Umständen effizient – und somit ohne Verbesserungspotential – erfolgt sei.

Folgende Rahmenbedingungen beeinflussen den Transformationsprozess und stehen nicht unter der Kontrolle der Entscheidungsträger der Firmen:

²⁵Dieser Nachteil ist nicht eine Eigenheit der nichtparametrischen Ansätze und tritt auch bei der *Stochastic Frontier* und *Cost Function Analysis* auf.

1. Gesetzliche Restriktionen und Regulierung seitens der Behörden, die grundsätzlich die Entscheidungsfreiheit der Manager beschränken aber nur einen Teil der Firmen betreffen, sind eine Ursache für unterschiedliche Rahmenbedingungen. Solche Unterschiede existieren besonders bei interregionalen bzw. internationalen Effizienzvergleichen. Absorbieren z.B. Umweltschutzmassnahmen Ressourcen, die als Produktionsfaktoren in eine Effizienzmessung eingehen, werden Firmen, die strengere Umweltschutzaufgaben zu erfüllen haben, *ceteris paribus* als ineffizienter deklariert.
2. Standortfaktoren. Eine Ursache für standortabhängige Einflussfaktoren ist in der regional unterschiedlichen Qualität der Inputs und Outputs zu suchen. Zu erwähnen ist beispielsweise das Ausbildungsniveau der Arbeitskräfte in einer Region²⁶ oder klimatische Einflussfaktoren.
3. Unterschiedliche Anreizstrukturen, z.B. infolge der Rechtsform (privatwirtschaftlich bzw. öffentlich-rechtlich), führen zu unterschiedlichen Zielfunktionen der Entscheidungsträger. So dürften etwa die Manager gewinnorientierter Unternehmungen andere Ziele verfolgen als die Entscheidungsträger in genossenschaftlich organisierten Betrieben. Bei der Effizienzmessung wird aber die Maximierung der Effizienz explizit in der Zielfunktion unterstellt.

Es ist wichtig, dass diese den Transformationsprozess beeinflussenden Faktoren in der Effizienzmessung berücksichtigt werden: Grundsätzlich sind zwei Vorgehensweisen denkbar. Einerseits können die Einflussfaktoren in die Effizienzmessung integriert werden. Die geschätzten Effizienzwerte sind dann direkt als solche zu interpretieren, da für die Einflussfaktoren kontrolliert wurde. Dieser Ansatz wird als *One Stage Approach* bezeichnet. Andererseits kann die relative Distanz zur effizienten Grenze ungeachtet der Einflussfaktoren in einem ersten Schritt geschätzt werden. Das Resultat wird anschliessend in einem zweiten Schritt in die Komponenten Effizienz und Effekte der übrigen Einflussfaktoren zerlegt. Entsprechend wird diese Vorgehensweise als *Two Stage Approach* bezeichnet. Beide Ansätze werden in den Abschnitten 2.3.1 und 2.3.2 dargestellt.

Folgerung 2.2 *Die relative Distanz zur effizienten Grenze kann in der Regel nicht direkt*

²⁶Gehen die bewerteten Inputs und Outputs in die Effizienzmessung ein, sollten die Qualitätsunterschiede – Wettbewerb auf den Faktor- und Gütermärkten vorausgesetzt – über die Preise kompensiert werden und keine Konsequenzen für die Effizienz haben. Im Allgemeinen ist eine solche Bewertung *a priori* nicht möglich und erfolgt erst im Rahmen der Effizienzmessung.

als Effizienz interpretiert werden. Die Rahmenbedingungen, die den Transformationsprozess beeinflussen aber trotzdem nicht Produktionsfaktoren im eigentlichen Sinne darstellen, müssen in die Analyse der Effizienz eingehen bzw. adäquat berücksichtigt werden.

2.3.1 Die One-Stage-Ansätze

Sollen verschiedene Gruppen von Firmen miteinander verglichen werden (z.B. privatwirtschaftlich vs. öffentlich-rechtlich organisierte), haben Charnes, Cooper und Rhodes (1981) vorgeschlagen, dass jeweils eine DEA für beide Gruppen separat durchgeführt wird. Die Beobachtungen werden dann auf die jeweilige GPM projiziert, d.h. sie werden um ihre Ineffizienz bereinigt. Schliesslich wird eine DEA mit allen projizierten Punkten durchgeführt und die so ermittelte Effizienz ist die Grundlage für die statistische Auswertung. Die Effizienzdifferenzen der beiden Gruppen sind dann ein Indiz für die Wirkung der Gruppenzugehörigkeit. Die Trennschärfe der Differenz kann mit einem Mittelwerttest (T-Test) oder Rangsummentest (Mann-Whitney-U-Test) bestimmt werden.

Eine weitere Methode stammt von Banker und Morey (1986). Die Idee ist, dass verschiedene Gruppen mit unterschiedlich wünschenswerten Rahmenbedingungen auch an verschiedenen Referenzsets gemessen werden. Vorausgesetzt wird dabei, dass die Rahmenbedingungen bezüglich ihres Einflusses auf den Transformationsprozess geordnet werden können. Als Beispiel wird der Standort von Bankfilialen herangezogen. Es wird unterstellt, dass 1) Bankfilialen in Stadtzentren gegenüber den 2) Bankfilialen in der Agglomeration und diese wiederum gegenüber solchen in 3) landschaftlichen Regionen günstigere Rahmenbedingungen aufweisen. Die Autoren schlagen vor, dass die Beobachtungen mit den unvorteilhaftesten Rahmenbedingungen (Bankfilialen auf dem Land) an ihrem eigenen Referenzset gemessen werden, d.h. nur mit Filialen verglichen werden, die über dieselben Rahmenbedingungen verfügen. Die Resultate werden als Effizienz interpretiert. Das Referenzset der 2. Gruppe (Bankfilialen in der Agglomeration) beinhaltet alle Beobachtungen der 2. und 3. Gruppe. Die 1. Gruppe wird dann an allen Beobachtungen gemessen. Somit wird sichergestellt, dass die benachteiligten Bankfilialen nicht an der GPM der Beobachtungen mit den günstigeren Rahmenbedingungen evaluiert werden.

Ein gemeinsamer Nachteil beider Vorgehensweisen ist, dass sie auf der Unterteilung der Stichprobe in Teilstichproben basieren, was zur Folge hat, dass der relative und absolute Anteil der Beobachtungen, die auf der GPM liegen, zunimmt. Die Effizienzmessung verliert somit an Aussagekraft. Zudem wird verlangt, dass die Rahmenbedingung er-

fassende Variable über ein diskretes Messniveau verfügt. Der Vorschlag von Banker und Morey (1986) hat den zusätzlichen Nachteil, dass die Richtung der Wirkung der Rahmenbedingung *a priori* bekannt sein muss. Die Kenntnis der Wirkungsrichtung ist aber i.d.R. Gegenstand der Analyse und erst *a posteriori* bekannt.

Folgerung 2.3 *Die nichtparametrischen Methoden lassen die Integration der Rahmenbedingungen in die erste Stufe kaum befriedigend zu. Die mangelnde bzw. fast inexistente Umsetzung dieser Methoden in empirischen Studien bekräftigt die Vermutung, dass die nichtparametrischen One-Stage-Ansätze kaum von praktischer Relevanz sind.*

2.3.2 Die Two-Stage-Ansätze

Der *Two Stage Approach* wurde von Timmer (1971) zur Erklärung der Effizienzunterschiede in der landwirtschaftlichen Produktion zwischen den Gliedstaaten der USA verwendet. Beim *Two Stage Approach* wird in einem ersten Schritt die relative Distanz zur effizienten Grenze (anhand der DEA oder einer parametrischen Methode) geschätzt. Im zweiten Schritt wird die relative Distanz in einer Regressionsanalyse durch die Variablen, die die Rahmenbedingungen erfassen und den Transformationsprozess beeinflussen, erklärt. Folgende Vorteile ergeben sich durch diese Vorgehensweise. Erstens können die Effekte aller Einflussfaktoren gemeinsam erfasst werden (im Gegensatz zu den einstufigen Ansätzen, vgl. Abschnitt 2.3.1). Zweitens können stetige und diskrete Variablen zur Erklärung der Distanz verwendet werden. Drittens muss die Wirkungsrichtung der Rahmenbedingungen nicht *a priori* festgelegt werden; sie ist das Resultat der zweiten Stufe. Viertens können Hypothesentests zur Überprüfung der Trennschärfe der Zusammenhänge zwischen den erklärenden Variablen und der relativen Distanz durchgeführt werden. Diese Vorteile, kombiniert mit der relativ einfachen Handhabung, haben dem *Two Stage Approach* zur *Common Practice* verholfen.

Folgerung 2.4 *Der Two-Stage-Ansatz stellt ein geeignetes Mittel zur Integration unterschiedlicher Rahmenbedingungen dar. Verschiedene Ursachen für heterogene Rahmenbedingungen bei der Produktion können gemeinsam berücksichtigt werden, wobei die Anforderungen an das Messniveau der Variablen nicht eingeschränkt ist.*

Trotzdem blieb auch der *Two Stage Approach* nicht unkritisiert. Vier Vorbehalte werden in der Literatur erwähnt [vgl. z.B. Lovell (1993), S. 55 ff. oder Coelli et al. (1998), S. 171]:

1. Die Effizienzwerte sind auf Werte zwischen null und eins beschränkt. Die übliche Normalverteilungsannahme bei Regressionen ist somit i.d.R. nicht erfüllt.
2. Die Unterscheidung zwischen Inputs und Outputs einerseits und erklärenden Variablen andererseits ist nicht immer eindeutig und muss mit äusserster Vorsicht erfolgen.
3. Eine eventuelle Korrelation zwischen Input- und Outputvariablen mit den erklärenden Variablen der zweiten Stufe wird als mögliche Ursache für verzerrte Resultate erwähnt.
4. Es wird nur die radiale Effizienz berücksichtigt. Die Slacks, eine nicht zu vernachlässigende Form der Ineffizienz, werden von der Analyse ausgeschlossen.

Im Folgenden werden diese vier Punkte in Bezug auf den nichtparametrisch-ökonometrischen und den rein ökonometrischen *Two-Stage*-Ansatz überprüft. Punkt 1 betrifft beide Ansätze gleichermassen. Das Problem ist einfach zu handhaben, indem die geschätzten Effizienzwerte transformiert werden. Alternativ bieten sich Methoden der *Limited Dependent Variable Regression* an, welche sich in der Praxis gegenüber der Transformation der Effizienzwerte durchgesetzt haben.²⁷

Der zweite Kritikpunkt ist sowohl beim nichtparametrisch-ökonometrischen wie auch beim rein ökonometrischen *Two-Stage*-Ansatz relevant. Grundsätzlich wird diese Unterteilung durch den Transformationsprozess definiert. Allerdings ist die Unterteilung nicht immer einfach und eindeutig. Ray (1988), der als erster den nichtparametrisch-ökonometrischen Ansatz präsentierte, nimmt zu diesem Punkt Stellung. Er unterteilt die den Produktionsprozess definierenden und beeinflussenden Variablen in 1) diskretionäre und 2) nichtdiskretionäre Variablen. Die diskretionären Variablen, die durch die Entscheidungsträger kontrolliert werden können, sollen dabei in die erste Stufe einfließen. Die nichtdiskretionären Variablen sollen, da diese mindestens in der kurzen Frist fixiert sind, in der zweiten Stufe die Effizienz erklären. Der Autor weist darauf hin, dass die Unterscheidung in diskretionäre und nichtdiskretionäre Variablen nicht *a priori* gegeben ist und folglich dem Forschungsziel entsprechend bestimmt werden soll.

²⁷Neben dem beschränkten Wertebereich der Effizienzwerte ergibt sich in empirischen Anwendungen oft eine Häufung am oberen Ende des Wertebereichs (*Ceiling Effect*), welche auch durch eine Transformation nicht eliminiert werden kann.

Lovell (1993) erläutert den dritten Punkt anhand des rein ökonometrischen *Two-Stage*-Ansatzes. Sind die Variablen der ersten Stufe (die Regressoren der Produktion bzw. der Kosten) mit denjenigen der zweiten Stufe (den Regressoren der Effizienz) korreliert, sind die Resultate der ersten Stufe verzerrt, da ein *Omitted Variable Problem* vorliegt. Dieses *Omitted Variable Problem* kann umgangen werden, indem die Variablen der zweiten Stufe in die Schätzung der Effizienz integriert, d.h. in einem ökonometrischen One-Stage-Modell zusammengefasst werden. Diese Vorgehensweise hat zur Folge, dass das *Omitted Variable Problem* durch ein Multikollinearitätsproblem ersetzt wird. Ohne weitere Erklärung folgt (Lovell, S. 54):

The principles involved in the two-stage versus the one-stage formulation apply equally well to noneconometric efficiency measurement models...

Die Problematik der zweistufigen, rein ökonometrischen Methode scheint eins zu eins auf den nichtparametrisch-ökonometrischen Fall übertragen worden zu sein. Als Referenz dienen in Lovell (1993) die Artikel von Olesen und Petersen (1991) und Ray (1988). In beiden Papieren findet sich aber keine Antwort auf die Frage, wieso korrelierte Input-Output-Variablen der DEA und Regressoren der zweiten Stufe zu verzerrt geschätzter Effizienz führen. McCarty und Yaisawarng (1993) fassen die beiden Kritikpunkte 2 und 3 zusammen (S. 285 ff.):

...the two-stage approach can be problematic when there is strong correlation between the first-stage inputs and the second-stage independent variables. If these variables were strongly correlated, then the claim that the two stages incorporate fundamentally different types of inputs, controllable and uncontrollable, becomes untenable. In this case, DEA scores computed in the first stage are likely to be biased in the sense that they would actually reflect effects of both categories of inputs.

Diese Argumentation basiert auf der Ansicht, dass eine starke Korrelation die Unterscheidung zwischen diskretionären und nichtdiskretionären Variablen erschwert oder mindestens widersprüchlich erscheinen lässt. Zwei Situationen sind, je nach Kausalität, denkbar. Die erste Möglichkeit ist, dass der Umfang der diskretionären Variablen effektiv durch die Entscheidungsträger bestimmt werden kann. Infolge der Korrelation ist dann aber die nichtdiskretionäre Variable der zweiten Stufe nicht unabhängig: Eine Veränderung müsste zu einer entsprechenden Veränderung der nichtdiskretionären Variablen führen. Die zweite

Möglichkeit ist, dass die nichtdiskretionären Variablen effektiv fixiert sind, was implizieren würde, dass auch die diskretionären Variablen nicht durch die Entscheidungsträger beeinflusst werden können. Die Unterteilung in diskretionäre bzw. nichtdiskretionäre Variablen ist dann aber widersprüchlich. Die geschätzte Effizienz ist dann – so die Folgerung von McCarty und Yaisawarng (1993) – verzerrt, da sie die Effekte beider Variablenkategorien widerspiegeln, womit wieder das *Omitted Variable Problem* angesprochen wird. In Coelli et al. (1998) wird in Bezug auf den dritten Kritikpunkt auf Seite 171 folgende Aussage gemacht:

One disadvantage of the two-stage method is that if the variables used in the first stage are highly correlated with the second-stage variables then the results are likely to be biased.

Coelli et al. (1998) verzichten darauf, zu erklären, wieso die Resultate verzerrt sind. Leider findet sich ebenda auch keine entsprechende Literaturangabe, die diese offenen Fragen zu klären vermag. Die Vermutung ist berechtigt, dass der Erklärungsansatz des rein ökonometrischen *Two-Stage*-Ansatzes auf den nichtparametrisch-ökonometrischen Ansatz übertragen wurde. Diese Vermutung wird auch durch ein Mail von Tim Coelli vom 7. Juni 2001 bekräftigt. Auf die Frage, wie er seine Aussage begründe, antwortet er:

I have seen this stated in various papers – I base my comments on the idea of omitted variable bias in regression – the effect should be similar.

Dem vierten Kritikpunkt ist anzufügen, dass entweder ein Effizienzmass verwendet wird, das strikt monoton ist und eine eindeutige Effizienzindikation erlaubt,²⁸ oder das Ausmass der *Slacks* neben der Effizienz in der zweiten Stufe simultan in einem *Seemingly-Unrelated-Regression*-Modell zu schätzen [vgl. Fried, Lovell und Vanden Eeckaut (1993)].

Folgerung 2.5 *Zusammenfassend ist festzustellen, dass die in der Literatur aufgeführten Bedenken gegenüber dem nichtparametrisch-ökonometrischen Two-Stage-Ansatz sich auf den Omitted Variable Bias infolge korrelierter Variablen der ersten und zweiten Stufe reduzieren lässt. Eine theoretische Grundlage dieser Argumentation wurde aber für den nichtparametrisch-ökonometrischen Ansatz bislang nicht erbracht.*

²⁸Als Beispiele sind die *Cone-Ratio*-DEA Versionen [vgl. Charnes et al. (1990) und Thompson et al. (1990), bzw. Abschnitt 2.2.3] und die Additiven-DEA-Versionen zu nennen [vgl. z.B. Cooper, Seiford und Tone (2000)].

Kapitel 3

Konsistenzprobleme der Data Envelope Analysis

In der empirischen Forschung interessieren neben der Effizienz der analysierten Produktionseinheiten oft die Determinanten der Effizienz. Aus diesem Grund, und weil die DEA die Integration der Rahmenbedingungen kaum befriedigend zulässt (vgl. Abschnitt 2.3.1), wird üblicherweise die relative Distanz zur GPM statistisch bzw. ökonometrisch ausgewertet. Anhand von Hypothesentests im Rahmen des *Two-Stage*-Ansatzes können Rückschlüsse über die Wirkungsweise und -richtung der Rahmenbedingungen und der Faktoren, die den Produktionsprozess beeinflussen, gemacht werden. Solche Hypothesentests bedingen allerdings, dass die Schätzer der Effizienz, d.h. die durch die DEA ausgewiesene relative Distanz zur GPM, gewissen statistischen Qualitätskriterien genügen. Sind diese Kriterien nicht erfüllt, resultieren unter Umständen Fehler bei den Hypothesentests. In diesem Kapitel wird untersucht, unter welchen Umständen Hypothesentests im Rahmen des *Two-Stage*-Ansatzes zu falschen bzw. korrekten Aussagen führen. Im Zentrum der Fragestellung steht der Vergleich der durchschnittlichen Effizienz zweier Gruppen von Beobachtungen.

Der folgende Gedanke bildet die Ausgangslage zu diesem Kapitel: Monotone Effizienzmasse haben die Eigenschaft, dass eine Zunahme der Inputmengen bei konstanten Outputmengen oder eine Abnahme der Outputmengen bei konstanten Inputmengen in einer Abnahme der Effizienz resultiert. In Abschnitt 2.2 wurde erläutert, dass das Effizienzmass der DEA schwach monoton ist: Werden *alle* Inputs einer Beobachtung um einen bestimmten Faktor erhöht, sinkt die Effizienz dieser Beobachtung um den reziproken

Wert des entsprechenden Faktors. Ein nahe liegender Schluss ist, dass eine proportionale Ausdehnung (Reduktion) aller Inputs (Outputs) bei konstanten Outputmengen (Inputmengen) bei einer Untergruppe von Beobachtungen einen entsprechenden Einfluss auf die Effizienz dieser Beobachtungen und folglich auf die durchschnittliche Effizienz dieser Untergruppe haben muss. In diesem Kapitel wird nachgewiesen, dass dies nur unter gewissen Umständen gilt.

Der folgende Abschnitt 3.1 befasst sich mit der statistischen Fundierung der DEA. Im Abschnitt 3.2 wird für eine Transformationsfunktion mit einem Output und zwei Inputs grafisch gezeigt, unter welchen Bedingungen die Gefahr besteht, dass die Schätzer der Effizienz den statistischen Qualitätskriterien nicht genügen. Als problematisch erweist sich dabei, dass die Bedingungen, die die statistischen Qualitätskriterien der Schätzer gewährleisten, sehr restriktiv sind. In Abschnitt 3.3 werden die Erkenntnisse auf den allgemeinen Fall, d.h. mit beliebig vielen Inputs und Outputs, übertragen. Der letzte Abschnitt untersucht die Frage, wie in der empirischen Forschung überprüft werden kann, ob die Schätzer die Qualitätskriterien erfüllen oder nicht. Dazu schlage ich eine Kennzahl vor und es werden zwei Hypothesen bezüglich der Eigenschaften dieser Kennzahl formuliert, die anschliessend in Kapitel 4 überprüft werden.

3.1 Statistische Fundierung der DEA

Im Rahmen von Hypothesentests geht es darum, anhand von Schätzern auf Grund einer Stichprobe auf die Parameter der Grundgesamtheit zu schliessen. In diesem Kontext ist die DEA-Effizienz ein Schätzer für die Effizienz. Die folgende Diskussion bezieht sich nicht auf das Konzept der relativen Distanz zur GPM – sie bezieht sich auf die Eigenschaften der Schätzer der Effizienz, die anhand der DEA ermittelt werden.¹

Hypothesentests bedingen, dass die Schätzer bestimmte statistische Eigenschaften aufweisen: Die Schätzer sollten erwartungstreu und konsistent sein.² Erwartungstreue (Unverzerrtheit) bedingt, dass der Erwartungswert des Schätzers dem wahren Wert entspricht.

¹Die relative Distanz zur GPM ist mathematisch definiert und die ermittelten Effizienzwerte sind eindeutig bestimmt. Für diese Effizienzwerte existieren die unten definierten Begriffe Erwartungstreue und Konsistenz nicht. Erst wenn sie als Schätzer für die wahre Effizienz verwendet werden, sind die statistischen Eigenschaften der Erwartungstreue und der Konsistenz definiert.

²Eine weitere Eigenschaft der Schätzer, die nicht weiter behandelt wird, ist die Effizienz. Effizient ist eine Schätzmethode, wenn die Varianz der Verteilung des Schätzers bei gegebener Stichprobengrösse kleiner ist als bei allen alternativen Methoden.

Dies muss unabhängig von der Stichprobengrösse erfüllt sein. Unter dem Begriff der statistischen Konsistenz wird eine Eigenschaft verstanden, die verlangt, dass die geschätzten Parameter mit zunehmender Stichprobengrösse sich den wahren Werten annähern: Ein Schätzer ist konsistent, wenn er mit einer Wahrscheinlichkeit von eins gegen den wahren Wert konvergiert.³ Wenn die Schätzer unverzerrt und konsistent sind, gilt, dass unabhängig von der Stichprobengrösse die geschätzten Werte im Erwartungswert den wahren Werten entsprechen. Sind die Schätzer einer Methode verzerrt aber trotzdem konsistent, resultieren bei finiten Stichprobengrössen systematische Fehler.

Banker (1993) zeigt, dass die Schätzer der Effizienz bei der DEA verzerrt aber unter bestimmten Bedingungen konsistent sind.⁴ Neben Bedingungen, die bei der DEA infolge der Strukturannahmen der Technologiemenge *a priori* erfüllt sind (vgl. Kapitel 2), formuliert Banker eine Konsistenzbedingung bezüglich der Effizienzverteilung, wonach die Wahrscheinlichkeit einer fast vollständig effizienten Beobachtung in einer Grundgesamtheit grösser null sein muss.⁵ Diese Konsistenzbedingung wird im folgenden als die explizite Konsistenzbedingung bezeichnet. Je näher die effizientesten Beobachtungen in einer Stichprobe bei der wahren Grenze der Technologie liegen, desto dichter kommt auch die geschätzte Technologiemenge an die wahre Technologie heran. Die effizienten Beobachtungen haben die Aufgabe, die GPM auf das „wahre“ Niveau zu heben. Ist die Wahrscheinlichkeit einer fast vollständig effizienten Beobachtung grösser null, ist auch die Wahrscheinlichkeit, dass die geschätzte Technologiemenge das wahre Effizienzniveau erreicht, grösser null.

Die Schätzer der Effizienz sind verzerrt, falls die Wahrscheinlichkeit einer absolut effizienten Beobachtung in einer Stichprobe kleiner eins ist. Verzerrungen werden realisiert, da die effizientesten Beobachtungen, die nicht vollständig effizient sein müssen, bei der

³Eine Übersicht zur Konsistenz nichtparametrischer Methoden der Effizienzmessung bietet Grosskopf (1996). Die theoretische Grundlagen finden sich in Banker (1993), Banker (1996), Kneip et al. (1998) und Simar und Wilson (2000).

⁴Banker (1993) leitet zudem Bedingungen her, unter denen die geschätzte DEA-Effizienz *Maximum-Likelihood*-Schätzer sind. Dies ist der Fall, wenn a) die Dichtefunktion der Effizienzverteilung monoton zunehmend ist und b) eine Transformationsfunktion mit nur einem Output verwendet wird.

⁵Bei stetigen Dichtefunktionen ist die Wahrscheinlichkeit einer vollständig effizienten Beobachtung per Definition gleich null. Mit der Formulierung „fast vollständig effizient“ wird ein beliebig kleines Intervall unmittelbar unter dem oberen Ende der Effizienzverteilung definiert. Die Wahrscheinlichkeit einer fast vollständig effizienten Beobachtung, d.h. einer Beobachtung die in diesem Intervall liegt, ist dann strikt positiv, wenn die (stetige) Dichtefunktion am oberen Ende einen strikt positiven Wert annimmt.

DEA per Definition als effizient deklariert werden. Die Verzerrungen wirken folglich immer in dieselbe Richtung: Die Effizienz wird überschätzt. Sind die Anforderungen an die Konsistenz der DEA erfüllt (die Wahrscheinlichkeit einer fast vollständig effizienten Beobachtung ist grösser null), ist die Stichprobe aber so klein, dass im Erwartungswert keine effiziente Beobachtung auftritt, ist die geschätzte Effizienz im Erwartungswert verzerrt. Mit zunehmender Stichprobengrösse steigt die Wahrscheinlichkeit, dass eine effiziente Beobachtung auftritt und die erwartete Verzerrung nimmt ab.⁶ Die erwartete Verzerrung wirkt sich in zweierlei Hinsichten aus: Erstens weichen die geschätzten Effizienzwerte um den Betrag der erwarteten Verzerrung von den wahren Werten ab, und zweitens – als logische Konsequenz – verschiebt sich die geschätzte Effizienzverteilung in die entsprechende Richtung.

Wegen der mangelnden Erwartungstreue sind die Schätzer der Effizienz in empirischen Anwendungen der DEA immer verzerrt: Die Effizienz wird überschätzt. Für Hypothesentests bezüglich der Mittelwerte von zwei Gruppen von Beobachtungen ist aber nicht das Effizienzniveau, sondern die Differenz der durchschnittlichen Effizienz relevant. Ist die erwartete Verzerrung der Effizienz für alle Beobachtungen bzw. Gruppen von Beobachtungen identisch, sind die Effizienzdifferenzen unverzerrt, d.h. valide, auch wenn die Effizienz jeder einzelnen Beobachtung überschätzt wird.

Folgerung 3.1 *Im Erwartungswert sind die geschätzten Effizienzdifferenzen bereits dann unverzerrt und folglich valide, falls beide Gruppen eine identische erwartete Verzerrung aufweisen (sie muss nicht gleich null sein). Sind die erwarteten Verzerrungen nicht gleich gross, sind die Effizienzdifferenzen nicht valide und es besteht die Gefahr, dass dieser Fehler sich auf die Ebene der Hypothesentests überträgt und zu falschen Rückschlüssen führt.⁷*

Eine weitere, implizite Annahme ist, dass die Beobachtungen unabhängig und identisch in der Technologiemenge verteilt sind. Diese Annahme wird sich als problematisch erweisen, da sie – wie später argumentiert wird – kaum mit ökonomischem Handeln der

⁶Als erwartete Verzerrung wird im Folgenden die Differenz zwischen der geschätzten erwarteten Effizienz und der wahren Effizienz bezeichnet (vgl. Kennedy (1998), S. 24).

⁷Die Problematik der Erwartungstreue und der Konsistenz stellt sich nicht nur im Rahmen des *Two-Stage*-Ansatzes, sie ist direkt an die DEA-Schätzer der Effizienz geknüpft. Diese statistischen Eigenschaften charakterisieren die Güte der Schätzer, die in der zweiten Stufe zur Anwendung kommen. Somit wird die Qualität der Schätzer (die erste Stufe) auf die Hypothesentests (zweite Stufe) übertragen.

Produzenten in Einklang zu bringen ist. Die Konsequenz der Verletzung dieser impliziten Annahme ist unter Umständen, dass die Schätzer der Effizienz für einzelne Beobachtungen oder Teilstichproben nicht konsistent bzw. *unterschiedlich stark* verzerrt und somit nicht valide sind.⁸

Ausgehend von einer *idealen* Konstellation werden mögliche Ursachen für unterschiedlich stark verzerrt geschätzte Effizienz bzw. für nicht valide Effizienzdifferenzen gesucht. Da i.d.R. weder die wahre Effizienz noch der Erwartungswert der geschätzten Effizienz bekannt sind, können die Verzerrungen in einer empirischen Effizienzmessung auch nicht ermittelt werden. In diesem Kapitel wird daher wie folgt vorgegangen: Führt eine bekannte (und somit „wahre“) Veränderung der Input-Outputdaten einer Gruppe von Beobachtungen nicht zu einer entsprechenden Veränderung der geschätzten Effizienzwerte, besteht die Gefahr, dass die Effizienz nicht konsistent geschätzt werden kann. Da in einer solchen Situation auch die geschätzte Effizienzdifferenz unverändert bleiben wird, kann mit Sicherheit gesagt werden, dass diese nicht bzw. nur zufällig valide ist. Zieht dagegen eine Veränderung der Input-Outputdaten eine entsprechende Veränderung der geschätzten Effizienzdifferenz nach sich, ist die Effizienzdifferenz unverzerrt und somit valide.

3.2 Grafische Analyse

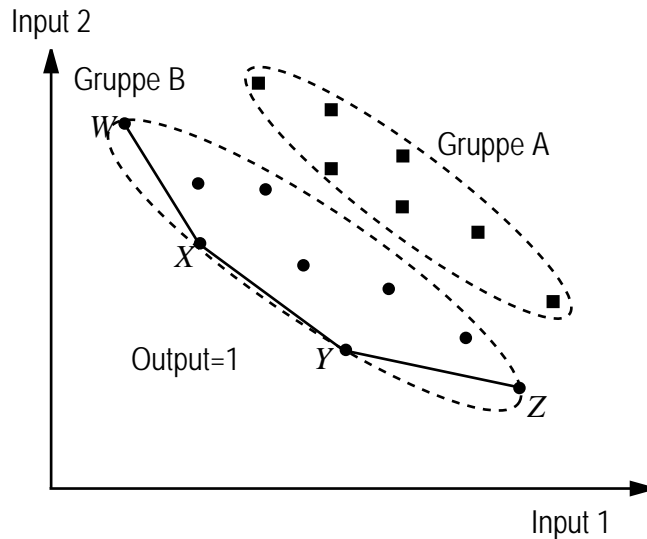
3.2.1 Die ideale Konstellation (die implizite Konsistenzbedingung ist erfüllt)

In diesem Abschnitt soll die *ideale* Konstellation erläutert werden, die in empirischen Anwendungen der DEA implizit unterstellt wird: Die Beobachtungen sind in Bezug auf die Input- und Outputverhältnisse unabhängig und identisch verteilt – die implizite Konsistenzbedingung ist erfüllt (vgl. Abbildung 3.1).⁹ In diesem Beispiel wird ein Output mit den zwei Produktionsfaktoren (Input 1 und 2) hergestellt; der Output ist auf eins normiert. Die GPM verläuft in Abbildung 3.1 durch die Punkte W , X , Y und Z . Dies ist die Ausgangslage nach der ersten Stufe, d.h. nach der DEA. In der zweiten Stufe

⁸Ob die Schätzer nicht konsistent oder unterschiedlich verzerrt sind hängt von den Annahmen bezüglich der Effizienzverteilungen ab.

⁹Die folgenden Ausführungen beschränken sich auf den dreidimensionalen Fall, d.h. die Summe der Anzahl Inputs und Outputs ist immer gleich drei. Der allgemein gültige Fall wird in Abschnitt 3.3 diskutiert.

Abbildung 3.1: Die ideale Konstellation mit unterschiedlich effizienten Gruppen



soll nun der Effekt der Gruppenzugehörigkeit, A oder B, auf die relative Distanz zur GPM werden.¹⁰ Die Input-Outputtransformationen der Gruppe A sind durch Quadrate gekennzeichnet und in der Ellipse rechts oben zusammengefasst, diejenigen der Gruppe B sind durch Punkte gekennzeichnet und in der Ellipse links unten zusammengefasst. Die Nullhypothese laute, dass beide Firmen gleich effizient produzieren.

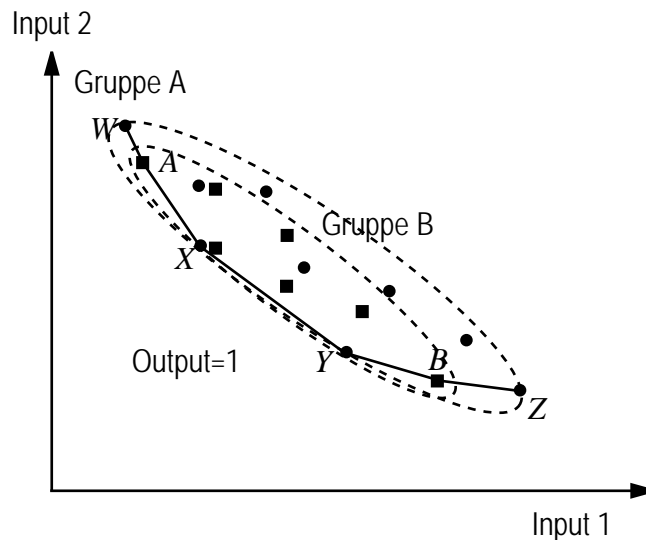
Offensichtlich fällt es den Firmen der Gruppe A schwerer, die beiden Inputs in den Output zu transformieren (für dieselbe Outputquantität werden mehr Inputs benötigt). Die effiziente Grenze wird ausschliesslich durch Firmen der Gruppe B definiert; die Zugehörigkeit zur Gruppe der effizienten Firmen korreliert mit der Gruppenzugehörigkeit. Gemäss Abbildung 3.1 scheint die Gruppenzugehörigkeit zudem mit der relativen Distanz zur effizienten Grenze zu korrelieren: Die beiden Gruppen weisen eine unterschiedliche durchschnittliche Effizienz auf. Ob diese Effizienzunterschiede statistisch signifikant sind, muss im Einzelfall geklärt werden. Da dieses Beispiel lediglich der Illustration dient, wird auf diese Auswertung verzichtet und davon ausgegangen, dass die Effizienzunterschiede signifikant sind. Folglich muss die Nullhypothese – beide Gruppen sind gleich effizient – verworfen werden.

In der Abbildung 3.2 ist eine weitere ideale Konstellation dargestellt. Nun sind die

¹⁰In diesem Beispiel sei die Gruppenzugehörigkeit die einzige Variable, die zu unterschiedlichen Rahmenbedingungen führt.

Effizienzunterschiede nicht mehr derart offensichtlich. Auch die Zugehörigkeit zur Gruppe der effizienten Firmen ist nicht mehr eindeutig: Sowohl Beobachtungen der Gruppe A wie auch der Gruppe B definieren die effiziente Grenze. Ob die Nullhypothese in diesem

Abbildung 3.2: Die ideale Konstellation mit gleich effizienten Gruppen



zweiten Beispiel verworfen werden muss, würde wieder Gegenstand einer statistischen Auswertung sein. In diesem illustrativen Beispiel wird davon ausgegangen, dass die Effizienzunterschiede nicht signifikant sind – die Nullhypothese kann nicht verworfen werden.

Ob die Schätzer der Effizienz verzerrt sind oder nicht, hängt in beiden Beispielen von der expliziten Konsistenzbedingung ab. Die implizite Konsistenzbedingung ist jedoch erfüllt und diese stellt sicher, dass für beide Gruppen dieselbe GPM als Referenz dient. Die Verzerrungen für beide Gruppen sind somit gleich gross und die Effizienzdifferenz ist unverzerrt und folglich valide (vgl. Folgerung 3.1). Die Aussagen der Hypothesentests¹¹ sind im Erwartungswert korrekt: Die Wahrscheinlichkeit eines Fehlers erster oder zweiter Art wird durch die Irrtumswahrscheinlichkeit der statistischen Tests korrekt repräsentiert.

Folgerung 3.2 *In der idealen Konstellation, die implizit unterstellt wird, sind die Effizienzdifferenzen im Erwartungswert unverzerrt und die Wahrscheinlichkeit eines Fehlers*

¹¹Die Nullhypothese besagt, dass beide Gruppen gleich effizient sind. Die Alternativhypothese lautet dagegen: Die Gruppen sind unterschiedlich effizient. Ein Fehler erster Art wird realisiert, wenn die Nullhypothese irrtümlich abgelehnt wird. Ein Fehler zweiter Art tritt dagegen auf, wenn die Nullhypothese irrtümlich nicht abgewiesen wird.

erster oder zweiter Art bei Hypothesentests wird adäquat durch die Irrtumswahrscheinlichkeit der statistischen Tests repräsentiert.

3.2.2 Die nicht-ideale Konstellation (die implizite Konsistenzbedingung ist verletzt)

In diesem Abschnitt wird eine Konstellation präsentiert, bei der die implizite Konsistenzbedingung verletzt ist, was i.d.R. zu nicht valid geschätzten Effizienzdifferenzen führt. Dies kann zur Folge haben, dass die Hypothesentests im Rahmen des *Two Stage Approaches* zu falschen Rückschlüssen führen. Wie im vorangegangenen Abschnitt wird zur Illustration ein extremes Beispiel gewählt, das in dieser Form selten auftreten dürfte und im Folgenden als die *nicht-ideale* Konstellation bezeichnet wird.

Die Leistung der Entscheidungsträger besteht aus drei Komponenten. Erstens sollen die innerbetrieblichen Abläufe effizient gestaltet, d.h. die technische Effizienz gesteigert werden. Zweitens muss die optimale Kombination von Inputs und Outputs unter Berücksichtigung der Rahmenbedingungen gewählt, d.h. die allokativen Effizienz erreicht werden. Drittens muss die optimale Betriebsgröße zur Erreichung der Skaleneffizienz realisiert werden. Die Ziele und auch die Massnahmen, die zur Erreichung dieser Ziele gewählt werden, sind das Resultat der Maximierung der Nutzenfunktion der Entscheidungsträger unter Nebenbedingungen. Neben den Präferenzen der Entscheidungsträger beeinflussen somit die Rahmenbedingungen das Kalkül der Entscheidungsträger. Diese Rahmenbedingungen umfassen, wie in Abschnitt 2.3 erläutert wurde,

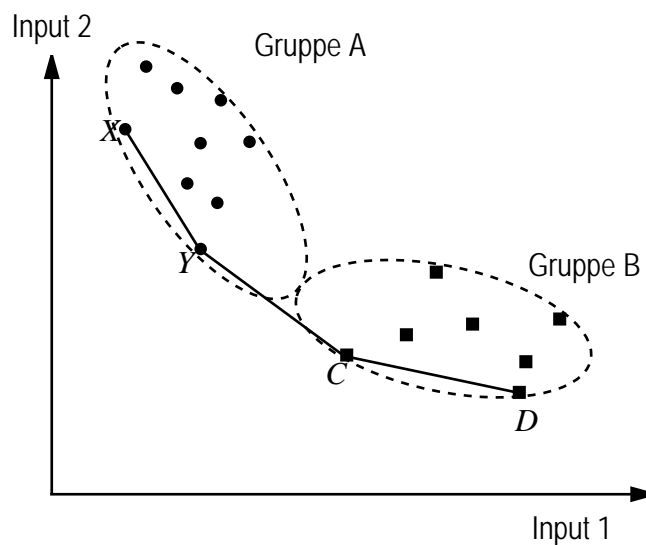
1. gesetzliche Restriktionen und Regulierung seitens der Behörden, die grundsätzlich die Entscheidungsfreiheit der Manager beschränken aber nur einen Teil der Firmen betreffen,
2. Standortfaktoren und
3. unterschiedliche Anreizstrukturen.

Eine Folge der unterschiedlichen Rahmenbedingungen sind einerseits unterschiedliche relative Preise und andererseits exogene Beschränkungen bei der Wahl der Inputs und Outputs. Falls sich die Rahmenbedingungen für die verschiedenen Gruppen von Entscheidungsträgern *systematisch* unterscheiden, werden diese auch systematisch unterschiedliche

Input- und Outputkombinationen wählen und folglich ist die implizite Konsistenzbedingung verletzt.

In Abbildung 3.3 ist eine Konstellation mit zwei Gruppen, die mit unterschiedlichen Rahmenbedingungen konfrontiert sind, dargestellt. Die Firmen der Gruppe A produzieren im Vergleich zu den Beobachtungen der Gruppe B mit einem grösseren Anteil des zweiten Inputs. Die Konsequenz ist, dass sich die Technologien der beiden Gruppen (d.h. die rele-

Abbildung 3.3: Die nicht-ideale Konstellation



vanten Teilmengen der Technologiemenge) unterscheiden; die beiden Technologien haben – wie im Extremfall in Abbildung 3.3 – keine Schnittmenge. Der Abschnitt YC der effizienten Grenze ist dann lediglich eine Verbindung ohne empirische Relevanz, die Konsequenz der Konvexitätsannahme der Technologiemenge. Alle Input-Outputtransformationen der Firmen der Gruppe A lassen sich durch die Punkte X und Y abbilden, während diejenigen der Gruppe B durch die Punkte C und D abgebildet werden. In einer solchen Konstellation sind die geschätzten Effizienzwerte einer DEA mit den Beobachtungen beider Gruppen mit den Effizienzwerten der gruppenweise getrennt durchgeführten DEA identisch.

Folgerung 3.3 *Werden nutzenmaximierende Entscheidungsträger unterstellt und sind die Rahmenbedingungen nicht identisch, werden die Entscheidungsträger unterschiedliche, von den Rahmenbedingungen abhängige Input- und Outputkombinationen wählen. Die Technologien (als Teilmenge der Technologiemenge) der einzelnen Gruppen unter-*

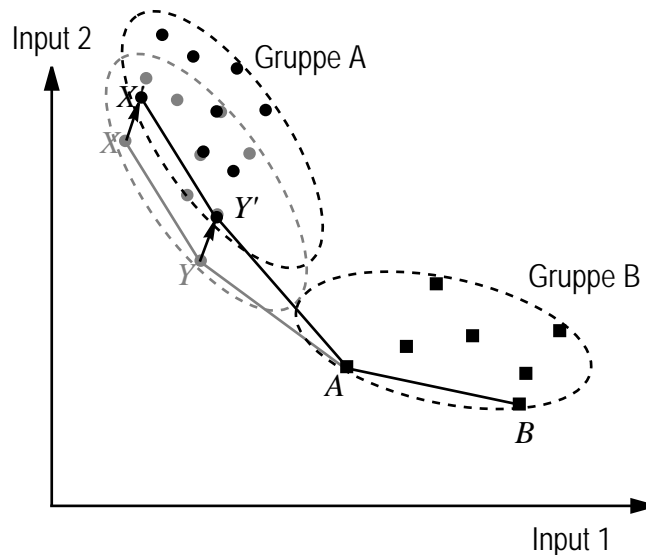
scheiden sich in der Folge. Im Extremfall weisen die Technologien keine Schnittmenge auf: Die Gruppen definieren gruppenspezifische, d.h. voneinander unabhängige, effiziente Abschnitte auf der GPM. Dies widerspricht der impliziten Konsistenzbedingung, wonach die Beobachtungen unabhängig und identisch in der Technologiemenge verteilt sein müssen.

Wird in einer nicht-idealen Konstellation die Zugehörigkeit zur Gruppe der effizienten Beobachtungen und das Ausmass der Effizienz statistisch ausgewertet, sind folgende Punkte zu beachten: Zum einen ist die Zugehörigkeit zur Gruppe der effizienten Firmen unabhängig von den Rahmenbedingungen, da A und B separate und somit unabhängige Teilmengen der Technologiemenge definieren. Folglich kann in einer solchen Konstellation die Zugehörigkeit zu einer Gruppe (A oder B) nicht die Zugehörigkeit zur Gruppe der effizienten Firmen erklären.¹² Zum anderen lässt die Auswertung der Distanz zur GPM keinen Rückschluss auf die Wirkung der gruppenspezifischen Rahmenbedingungen auf die Effizienzunterschiede zu. Gruppenspezifische Ineffizienz kann nicht erkannt werden. Dieser Sachverhalt ist in Abbildung 3.4 illustriert. Allen Input-Outputtransformationen der Gruppe A wird radiale gruppenspezifische Ineffizienz zugeführt, d.h. die Inputvektoren der Beobachtungen der Gruppe A werden mit einem einheitlichen Skalar multipliziert. Die ursprünglichen Input-Outputtransformationen der Gruppe A der Abbildung 3.3 sind grau dargestellt. Durch die zusätzliche gruppenspezifische Ineffizienz verschieben sich die neuen Input-Outputtransformationen nach rechts oben; mit ihnen verschiebt sich auch der effiziente Abschnitt der Gruppe A *pari passu* (vgl. Abschnitt $X'Y'$ der Abbildung 3.4). Die relative Distanz zur GPM bleibt unverändert und die gruppenspezifische Ineffizienz kann nicht erkannt werden. Offensichtlich ist dies ein Widerspruch, da die Beobachtungen der Gruppe A nun weniger effizient sein sollten: In diesem Beispiel sind die Effizienzdifferenzen nicht valide.

In der nicht-idealen Konstellation ist die implizite Konsistenzbedingung verletzt und es ist nicht sichergestellt, dass für beide Gruppen eine gemeinsame GPM als Referenz dient.

¹²Zu beachten ist, dass auch die relative Gruppengrösse einen Einfluss auf die Erklärungskraft der erklärenden Variable haben kann. Je kleiner eine Gruppe in Relation zur Stichprobe ist, desto grösser ist – in der nicht-idealen Konstellation – *ceteris paribus* der Anteil der effizienten Beobachtungen dieser Gruppe. Umfasst die Gruppe A im Extremfall nur die Beobachtungen X und Y (beide definieren in Abbildung 3.3 zusammen einen effizienten Abschnitt), sind 100 Prozent der entsprechenden Gruppe effizient, was grundsätzlich zu einer höheren Trennschärfe bei den statistischen Tests führen sollte. In einer solchen Situation müsste die Interpretation der statistischen Tests mit Vorsicht erfolgen. Die Resultate der statistischen Tests wären dann eher auf die Grösse der Teilstichprobe der Gruppe A und den unabhängigen effizienten Abschnitt zurückzuführen.

Abbildung 3.4: Gruppenspezifische Ineffizienz



Es ist somit nicht garantiert, dass die Verzerrungen der Schätzer der Effizienz infolge der fehlenden Erwartungstreue der DEA für beide Gruppen gleich gross ist. Die Effizienzdifferenzen sind folglich nicht bzw. bestenfalls zufällig unverzerrt und valide. Wieder laute die Nullhypothese, dass beide Gruppen gleich effizient die beiden Inputs in den Output transformieren. Eine statistische Auswertung der Konstellation in Abbildung 3.4 ergäbe aber, dass die Nullhypothese nicht verworfen werden kann, obwohl tatsächlich gruppenspezifische Ineffizienz vorliegt: Es liegt ein Fehler zweiter Art vor, die Nullhypothese wird zu unrecht nicht verworfen.

Folgerung 3.4 *In der nicht-idealen Konstellation können Effizienzdifferenzen zwischen Gruppen von Beobachtungen, die auf die gruppenspezifischen Rahmenbedingungen zurückzuführen sind, nicht erkannt werden. Es besteht die Gefahr nicht valider Effizienzdifferenzen, diese sind nicht bzw. bestenfalls zufällig unverzerrt. Die Wahrscheinlichkeit eines Fehlers zweiter Art nimmt folglich zu.*

Die Konsistenzbedingungen sind bei der DEA an zwei Verteilungsannahmen geknüpft. Sind diese erfüllt, sind die Effizienzwerte in kleinen Stichproben zwar verzerrt, die Effizienzdifferenzen sind im Erwartungswert aber unverzerrt. Die hier erfolgte Unterteilung in die ideale bzw. nicht-ideale Konstellation dient lediglich der Illustration. Trotzdem soll

die Unterteilung der GPM in ideale und nicht-ideale Konstellationen nun erweitert und präzisiert werden. Dies ist sinnvoll, weil die Konsequenzen der verletzten Konsistenzbedingungen innerhalb jeder Kategorie homogen sind.

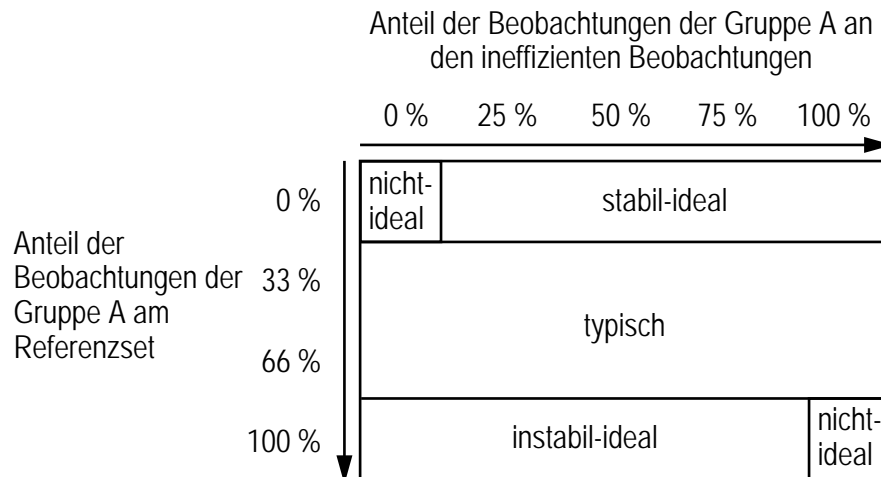
Die GPM besteht einerseits i.d.R. aus mehreren Abschnitten und es ist sehr wahrscheinlich, dass sich in der Praxis die effiziente Grenze sowohl aus idealen wie auch aus nicht-idealen Abschnitten zusammensetzt. Andererseits wurde bisher nur zwischen den beiden Extremen, der idealen (implizite Konsistenzbedingung erfüllt) und der nicht-idealen Konstellation (implizite Konsistenzbedingung nicht erfüllt), unterschieden, in der empirischen Forschung dürften diese extremen Konstellationen aber kaum in reiner Form auftreten. Denn auch die Frage, ob die implizite Konsistenzbedingung erfüllt ist kann nicht mit ja oder nein beantwortet werden. Eine eindeutige Antwort zu dieser Frage existiert lediglich im Zusammenhang mit einem bestimmten Signifikanzniveau.

Zur Klassifizierung eines Abschnittes der GPM müssen zwei Punkte beachtet werden. Einerseits ist die Zusammensetzung des Referenzsets und andererseits die Zusammensetzung der dazugehörenden, d.h. auf das entsprechende Referenzset projizierten, ineffizienten Beobachtungen in Bezug auf die Gruppenzugehörigkeit relevant. Diese Klassifizierung soll hier anhand eines einfachen Beispiels mit insgesamt sieben Beobachtungen veranschaulicht werden. Das Referenzset besteht aus drei Beobachtungen, auf das vier ineffiziente Beobachtungen projiziert werden. Wird das Referenzset ausschliesslich durch Beobachtungen der Gruppe A¹³ definiert, ist der „Anteil der Beobachtungen der Gruppe A am Referenzset“ gleich 100% (vgl. vertikaler Pfeil in Abbildung 3.5). Da annahmegemäss vier Beobachtungen ineffizient sind, kann der „Anteil der ineffizienten Beobachtungen der Gruppe A“ an den ineffizienten Beobachtungen die Werte 0, 25, 50, 75 und 100% annehmen (vgl. horizontaler Pfeil).

Die nicht-ideale Konstellation liegt dann vor, wenn beide Anteile die Werte 0 bzw. 100% annehmen. Sind beide Werte gleich 100, definiert die Gruppe A den effizienten Abschnitt und alle auf diesen Abschnitt projizierten Beobachtungen entstammen ebenfalls der Gruppe A (vgl. Abschnitt *XY* der GPM in Abbildung 3.3). Die Inputs der Beobachtungen der Gruppe A können variiert werden, ohne dass dies Konsequenzen auf die Effizienz hätte. Sind dagegen beide Werte gleich null, definieren die Beobachtungen der Gruppe B das Referenzset und alle ineffizienten Beobachtungen entstammen ebenfalls dieser Gruppe. Auch diese Konstellation ist nicht-ideal (vgl. Abschnitt *CD* der GPM in Abbildung 3.3).

¹³Die sieben Beobachtungen sind nicht fix auf die beiden Gruppen aufgeteilt.

Abbildung 3.5: Klassifizierung der Abschnitte



Die ideale Konstellation tritt in zwei Varianten auf. Eine ideale Konstellation ist *stabil*, wenn das Referenzset ausschliesslich von Beobachtungen der Gruppe B definiert wird und somit nicht von der gruppenspezifischen Ineffizienz betroffen ist und zudem mindestens eine ineffiziente Beobachtung der Gruppe A angehört (vgl. Abschnitte *WX*, *XY* und *YZ* der GPM in Abbildung 3.1). Stabil ist diese Konstellation deshalb, weil zusätzliche gruppenspezifische Ineffizienz die Beobachtungen der Gruppe A immer weiter von der effizienten Grenze entfernt.

Instabil-ideal ist dementsprechend eine Konstellation, wenn das Referenzset ausschliesslich aus Beobachtungen der Gruppe A definiert wird, mindestens eine ineffiziente Beobachtung dagegen der Gruppe B entstammt. Instabil ist diese Situation, weil mit zunehmender gruppenspezifischer Ineffizienz bei der Gruppe A der effiziente Abschnitt nach und nach durch die ursprünglich ineffizienten Beobachtungen der Gruppe B definiert wird. Ein instabil-idealer Abschnitt wird mit zunehmender gruppenspezifischer Ineffizienz stabil-ideal.¹⁴

Die vierte Möglichkeit ist die *typische* Konstellation, bei der das Referenzset aus Beobachtungen beider Gruppen besteht (vgl. Abschnitte *WA*, *AX*, *YB* und *BZ* der GPM in Abbildung 3.2).

¹⁴Die Unterteilung in stabil-ideal bzw. instabil-ideal ist von der Richtung der Veränderung der Daten abhängig. Werden die Inputs proportional reduziert (statt wie im Beispiel erhöht), wird die stabil-ideale Konstellation zur instabil-idealen Konstellation.

3.3 Mathematische Analyse

In diesem Abschnitt werden die Folgerungen 3.2 bis 3.4 mathematisch hergeleitet. Der Vorteil der mathematischen Analyse im Vergleich zum grafischen Ansatz besteht darin, dass die Anzahl der Inputs und Outputs nicht auf Werte von drei beschränkt ist. Es soll die Frage geklärt werden, ob und wie die Zielfunktionen der Linearen Programme 2.15 bzw. 2.19 auf die gruppenspezifische Ineffizienz reagieren. Ist die Zielfunktion in der nicht-idealen Konstellation – wie von Folgerung 3.4 postuliert – invariant in Bezug auf gruppenspezifische Ineffizienz oder nimmt sie mit zunehmender gruppenspezifischer Ineffizienz ab (wie in der idealen Konstellation)?

In der folgenden Gleichung 3.1 ist das Lineare Programm für die inputorientierte Version der DEA mit konstanten Skalenerträgen explizit aufgeführt.

$$\begin{aligned}
 \max \quad & z_j = \nu^T Y_j \\
 \text{s.t.} \quad & \mu^T X_j = 1 \\
 & \nu^T Y_1 - \mu^T X_1 \leq 0 \\
 & \nu^T Y_2 - \mu^T X_2 \leq 0 \\
 & \nu^T Y_3 - \mu^T X_3 \leq 0 \\
 & \nu^T Y_4 - \mu^T X_4 \leq 0 \\
 & \nu^T Y_5 - \mu^T X_5 \leq 0 \\
 & \nu^T Y_6 - \mu^T X_6 \leq 0 \\
 & \nu^T, \mu^T \geq 0
 \end{aligned} \tag{3.1}$$

Jede der sechs¹⁵ Produktionseinheiten verwendet m Inputs, um damit s verschiedene Outputs zu erzeugen. X_j ist der $m \times 1$ dimensionale Inputvektor, Y_j der $s \times 1$ dimensionale Outputvektor der Beobachtung j . Entsprechend sind μ und ν die Vektoren der virtuellen Multiplikatoren mit derselben Dimension. Die Zielfunktion der Beobachtung j entspricht der Summe der Produkte der virtuellen Multiplikatoren mit den Outputs. Die erste Nebenbedingung restringiert die Summe der Produkte der virtuellen Multiplikatoren mit den Inputs auf den Wert eins. Jede weitere Nebenbedingung restringiert für jede der sechs Beobachtungen die Differenz der gewichteten Outputs und der gewichteten Inputs auf nicht-positive Werte (vgl. Abschnitt 2.2.3).

¹⁵Die Beschränkung der Anzahl Beobachtungen erfolgt einzig und allein aus Platzgründen. Theoretisch könnten dem Beispiel beliebig viele Beobachtungen hinzugefügt werden. Da die bestehenden sechs Beobachtungen in den folgenden drei Konstellationen alle möglichen Anordnungen erfassen, würden zusätzliche Beobachtung immer in der Nähe einer der sechs Beobachtungen liegen müssen.

Die Beobachtungen mit ungeraden Laufnummern (1, 3 und 5) gehören im Folgenden zur Gruppe A, die übrigen drei zur Gruppe B. Nun wird ein Faktor Ψ , der die gruppenspezifische Ineffizienz repräsentieren soll, in das Lineare Programm eingeführt. Wird für den Ineffizienzfaktor Ψ ein Wert von eins gewählt, ändert sich nichts. Ist Ψ dagegen grösser als eins, absorbieren die Produktionseinheiten entsprechend mehr Inputs und sind – bei gegebenem Outputvektor – folglich ineffizienter. Es wird angenommen, dass die ersten vier Beobachtungen effizient seien; die Nebenbedingungen 3.4, 3.5, 3.6 und 3.7 sind daher bindend.

$$\max \quad z_j = \nu^T Y_j \quad (3.2)$$

$$\text{s.t.} \quad \mu^T X_j = 1 \quad (3.3)$$

$$\nu^T Y_1 - \Psi^A \mu^T X_1 = 0 \quad (3.4)$$

$$\nu^T Y_2 - \Psi^B \mu^T X_2 = 0 \quad (3.5)$$

$$\nu^T Y_3 - \Psi^A \mu^T X_3 = 0 \quad (3.6)$$

$$\nu^T Y_4 - \Psi^B \mu^T X_4 = 0 \quad (3.7)$$

$$\nu^T Y_5 - \Psi^A \mu^T X_5 \leq 0 \quad (3.8)$$

$$\nu^T Y_6 - \Psi^B \mu^T X_6 \leq 0 \quad (3.9)$$

$$\nu^T, \mu^T \geq 0$$

Zur Überprüfung der Folgerung 3.4 wird eine Fallunterscheidung nötig: 1) Das Referenzset der zu evaluierenden Beobachtung j entstammt nicht derselben Gruppe (die *ideale* Konstellation, vgl. Abschnitt 3.3.1), 2) das Referenzset der zu evaluierenden Beobachtung entstammt derselben Gruppe (die *nicht-ideale* Konstellation, vgl. Abschnitt 3.3.2) und 3) einige Beobachtungen des Referenzsets entstammen derselben Gruppe, andere hingegen nicht (die *typische* Konstellation, vgl. Abschnitt 3.3.3). Die Vorgehensweise wird dabei in den folgenden drei Abschnitten immer dieselbe sein: Zuerst wird festgelegt, welche Beobachtungen effiziente Abschnitte definieren und welche Beobachtungen an diesen Abschnitten evaluiert werden sollen (diese Einteilung ergibt sich aus der Konstellation, die es zu analysieren gilt). Dann wird die gruppenspezifische, radiale Ineffizienz eingeführt: Im Vergleich zur Ausgangslage werden die Beobachtungen der Gruppe A ineffizienter gemacht. Schliesslich werden die Auswirkungen der Änderungen, d.h. der gruppenspezifischen Ineffizienz, analysiert. Zu diesem Zweck wird die Wirkungen der Änderung der Daten auf die Nebenbedingungen analysiert und dann deren Konsequenzen auf die Zielfunktion bestimmt. Dabei übernimmt die Nebenbedingung 3.3 eine zentrale Rolle, da sie

immer bindend sein *muss*: Hat die Veränderung der Daten einen Einfluss auf diese Nebenbedingung, muss diese durch eine reziproke Anpassung der virtuellen Multiplikatoren der Inputs ausgeglichen werden. Wirkt sich dies auf die übrigen Nebenbedingungen und somit auf die virtuellen Multiplikatoren der Outputs aus, zieht dies eine Veränderung der Zielfunktion nach sich. Wirkt sich diese Anpassung hingegen nicht auf die übrigen Nebenbedingungen aus, bleibt der Wert der Zielfunktion unverändert.

3.3.1 Mathematische Analyse der idealen Konstellation

In der stabil-idealen Konstellation besteht das Referenzset für die ineffiziente Beobachtung 5 (Gruppe A) ausschliesslich aus Beobachtungen der Gruppe B, d.h. die Beobachtungen 2 und 4 definieren das Referenzset. Die ineffiziente Beobachtung 6 der Gruppe B wird an der durch die Beobachtungen 1 und 3 (beide Gruppe A) definierten Grenze evaluiert (dies entspricht der instabil-idealen Konstellation). Nun wird der Gruppe A gruppenspezifische Ineffizienz zugeführt: $\Psi^A > 1$; für die Gruppe B gilt $\Psi^B = 1$.

Die stabil-ideale Konstellation

Der Ineffizienzfaktor Ψ^B ändert den Wert der Zielfunktion der beiden effizienten Beobachtungen 2 und 4 nicht, da der Ineffizienzfaktor Ψ^B gleich eins ist. Da die beiden Nebenbedingungen 3.4 und 3.6 der Beobachtungen 1 und 3, die durch das Ψ^A verändert werden, bei der Evaluation der effizienten Beobachtungen 2 und 4 irrelevant sind, gilt folglich¹⁶

$$z_j^{*\Psi} = z_j^* \quad j = 2, 4. \quad (3.10)$$

Die optimale Lösung der beiden effizienten Beobachtungen 2 und 4 nimmt in dieser Konstellation nach der Änderung der Daten denselben Wert an.¹⁷

Anders sieht die Situation für die ineffiziente Beobachtung 5 aus. Nach der Multiplikation der Inputvektoren mit dem Ineffizienzfaktor ist die linke Seite von Nebenbedingung 3.3 grösser eins. Da die Nebenbedingung 3.3 immer bindend ist, muss der ursprünglich optimale Vektor der virtuellen Multiplikatoren μ^* mit dem reziproken Wert von Ψ^A multipliziert werden. Nach wie vor besteht das Referenzset der Beobachtung 5 aus den Beobachtungen 2 und 4. Mit den um den reziproken Faktor von Ψ^A reduzierten virtuellen

¹⁶Wobei $z_j^{*\Psi}$ die optimale Lösung der Beobachtung j nach der Veränderung der Inputs ist.

¹⁷Dasselbe gilt für ineffiziente Beobachtungen (8, 10, ..., 20, ..., n) der Gruppe B, die auf diesen Abschnitt projiziert werden, aber in diesem Beispiel nicht enthalten sind.

Multiplikatoren nehmen die entsprechenden Nebenbedingungen 3.5 und 3.7 aber positive Werte an. Da diese bindend sein müssen, muss der Vektor der optimalen virtuellen Multiplikatoren der Outputs ν^* ebenfalls mit dem reziproken Wert von Ψ^A multipliziert werden. Dies wirkt sich aber auf die Zielfunktion der Beobachtung 5 aus, welche einen im Vergleich zur Ausgangslage (ohne gruppenspezifische Ineffizienz) kleineren Wert $z^{*\Psi}$ annimmt:

$$z_j^{*\Psi} = z_j^* \frac{1}{\Psi^A} \leq z_j^* \quad j = 5. \quad (3.11)$$

Die optimale Lösung der ineffizienten Beobachtung 5 sinkt um den reziproken Wert von Ψ^A . Diese Konstellation ist insofern als stabil zu bezeichnen, als dass das Referenzset nicht durch die gruppenspezifische Ineffizienz tangiert wird. Eine Zunahme der gruppenspezifischen Ineffizienz wirkt sich *immer* vollumfänglich auf die optimale Lösung z^* der Beobachtung 5 aus.

Folgerung 3.5 *In der stabil-idealen Konstellation ändert die gruppenspezifische Ineffizienz – unabhängig von deren Ausmass – den Wert der Zielfunktion um den reziproken Faktor der gruppenspezifischen Ineffizienz ($\frac{1}{\Psi^A}$). Die gruppenspezifische Ineffizienz wird vollumfänglich erkannt und folglich sind die Effizienzdifferenzen valide.*

Die instabil-ideale Konstellation

Wie wirkt sich die gruppenspezifische Ineffizienz auf die Zielfunktion der übrigen Beobachtungen der Gruppe A aus? Im Falle der beiden effizienten Beobachtungen 1 und 3 gilt, dass die proportionale Ausdehnung der Inputs wieder eine reziproke Reduktion der optimalen virtuellen Multiplikatoren der Inputs μ^* nach sich ziehen muss (sonst ist jeweils die Nebenbedingung 3.3 nicht erfüllt). Die relevanten Nebenbedingungen 3.4 und 3.6 sind nach wie vor automatisch erfüllt, da der Effekt der Ausweitung der Inputs und der Effekt der reziprok reduzierten virtuellen Multiplikatoren sich gegenseitig aufheben:

$$z_j^{*\Psi} = z_j^* \quad j = 1, 3. \quad (3.12)$$

Die optimalen virtuellen Multiplikatoren der Outputs bleiben unverändert und somit auch der Wert der Zielfunktion.¹⁸ Die gruppenspezifischen Ineffizienz führt in der instabil-idealen Konstellation zu einer Verschiebung des effizienten Abschnitts, der durch die Punkte 1 und 3 definiert wird, nach aussen, d.h. in Richtung mehr Inputs

¹⁸Dasselbe gilt auch für ineffiziente Beobachtungen (7, 9..., 19, ..., $n - 1$) der Gruppe A, die auf diesen Abschnitt projiziert werden, aber in diesem Beispiel nicht enthalten sind.

Schliesslich müssen noch die Konsequenzen der gruppenspezifischen Ineffizienz Ψ^A auf die Beobachtung Nr. 6 (der Gruppe B) analysiert werden. Mit der Verschiebung der effizienten Grenze wurden die optimalen virtuellen Multiplikatoren der Inputs μ^* des entsprechenden effizienten Abschnittes im Vergleich zur Ausgangslage um den reziproken Faktor Ψ^A reduziert. Da mit diesen reduzierten Multiplikatoren aber die Nebenbedingung 3.3 bei der Evaluation der Beobachtung 6 nicht erfüllt ist (sie nimmt den reziproken Wert von Ψ^A an), müssen diese wieder die ursprünglichen Werte annehmen, d.h. mit dem Faktor Ψ^A multipliziert werden. Die Folge ist dann aber, dass die relevanten Nebenbedingungen 3.4 und 3.6 nicht bindend sind und eine Erhöhung der virtuellen Multiplikatoren der Outputs ν^* um den Faktor Ψ^A möglich ist, was eine entsprechende Zunahme der Zielfunktion um den Faktor Ψ^A nach sich zieht:

$$z_j^{*\Psi} = z_j^* \Psi^A > z_j^* \quad j = 6. \quad (3.13)$$

Während das Referenzset durch die zusätzliche gruppenspezifische Ineffizienz nach aussen verschoben wird, ohne die optimale Lösung der Beobachtungen 1 und 3 zu tangieren, nimmt die optimale Lösung der Beobachtung 6 zu. Die Ineffizienzunterschiede nehmen folglich mit zunehmender gruppenspezifischer Ineffizienz ab. Wird die gruppenspezifische Ineffizienz der Gruppe A weiter erhöht, nimmt der Wert der Zielfunktion der Beobachtung 6 zu, bis er schliesslich den maximalen Wert von eins annimmt und gleichzeitig die Beobachtungen 1 oder 3 ineffizient werden.¹⁹ Diese Konstellation ist instabil. Der effiziente Abschnitt wird dann durch eine der beiden Beobachtungen 1 oder 3 und der Beobachtung 6 definiert: Die instabil-ideale Konstellation wird zur typischen Konstellation. Folglich wird eine weiter erhöhte gruppenspezifische Ineffizienz dieselben Konsequenzen haben wie in der weiter unten erläuterten typischen Konstellation (vgl. Abbildung 3.5).

Folgerung 3.6 *In der instabil-idealen Konstellation hat eine kleine Zunahme der gruppenspezifischen Ineffizienz keinen Einfluss auf die optimalen Lösungen z^* dieser Gruppe. Die Effizienz derjenigen Beobachtungen, die nicht dieser Gruppe angehören, nimmt jedoch proportional zur gruppenspezifischen Ineffizienz zu. Somit wirkt sich die gruppenspezifische Ineffizienz in vollem Ausmass auf die Effizienzunterschiede aus: Die gruppenspezifische Ineffizienz wird erkannt und folglich sind die Effizienzdifferenzen valide.*

Zu beachten ist, dass grössere Änderungen der gruppenspezifischen Ineffizienz dazu führen,

¹⁹Dies ist genau dann erreicht, wenn die gruppenspezifische Ineffizienz Ψ^A der Ineffizienz der am wenigsten ineffizienten Beobachtung der Gruppe B entspricht.

dass nach und nach ursprünglich effiziente Beobachtungen ineffizient werden. Das Referenzset besteht dann aber aus Beobachtungen beider Gruppen (die typische Konstellation). Bei einer weiteren Erhöhung der gruppenspezifischen Ineffizienz wird das Referenzset schliesslich vollständig durch Beobachtungen der Gruppe ohne gruppenspezifische Ineffizienz definiert (dies entspricht der stabil-idealen Konstellation).

3.3.2 Mathematische Analyse der nicht-idealen Konstellation

In der nicht-idealen Konstellation werden die ineffizienten Beobachtungen 5 und 6 an den effizienten Abschnitten evaluiert, die durch die effizienten Beobachtungen derselben Gruppe definiert werden: Die Beobachtung 5 wird an den Beobachtungen 1 und 3, die Beobachtung 6 dagegen an den Beobachtungen 2 und 4 evaluiert. Wieder wird gruppenspezifische radiale Ineffizienz bei der Gruppe A eingeführt: $\Psi^A > 1$, für die Gruppe B gilt wieder $\Psi^B = 1$. Für die Beobachtungen der Gruppe B ändert sich nichts, da die relevanten Inputmengen unverändert und folglich die relevanten Nebenbedingungen 3.5 und 3.7 wie in der Ausgangslage bindend bleiben:

$$z_j^{*\Psi} = z_j^* \quad j = 2, 4, 6. \quad (3.14)$$

Für die Beobachtungen der Gruppe A gilt, dass die jeweilige Erfüllung der Nebenbedingung 3.3 bedingt, dass die optimalen virtuellen Multiplikatoren der Inputs μ^* im Vergleich zur Ausgangslage um den reziproken Faktor Ψ^A reduziert werden müssen. Diese Veränderung der virtuellen Multiplikatoren gleicht aber die Ineffizienz in den Inputs in den relevanten Nebenbedingungen 3.4 und 3.6 aus. Damit bleiben diese bindend, ohne dass die virtuellen Multiplikatoren der Outputs ν^* angepasst werden müssen. Folglich ändert sich auch nicht der Wert der Zielfunktion:

$$z_j^{*\Psi} = z_j^* \quad j = 1, 3, 5. \quad (3.15)$$

In dieser Situation ist der Wert der Zielfunktion für beide Gruppen unabhängig von gruppenspezifischer Ineffizienz. Dies gilt solange, bis eine der beiden effizienten Beobachtungen 1 oder 3 von einer Beobachtung der Gruppe B oder einer konvexen Linearkombination von Beobachtungen dominiert wird. Erst dann reagiert die Zielfunktion auf eine weitere Erhöhung der gruppenspezifischen Ineffizienz.

Folgerung 3.7 *In der nicht-idealen Konstellation wirkt sich die gruppenspezifische Ineffizienz nie auf den Wert der Zielfunktion aus. Die Effizienzdifferenzen bleiben daher unverändert und folglich können diese nicht bzw. nur zufällig unverzerrt sein.*

3.3.3 Mathematische Analyse der typischen Konstellation

Nun soll die typische Konstellation analysiert werden: Das Referenzset der ineffizienten Beobachtungen (5 und 6) bilden Beobachtungen beider Gruppen. In diesem Beispiel bilden die beiden Beobachtungen 2 (Gruppe B) und 3 (Gruppe A) gemeinsam einen effizienten Abschnitt. Wieder wird die gruppenspezifische Ineffizienz bei der Gruppe A eingeführt: $\Psi^A > 1$ ($\Psi^B = 1$). Der relevante effiziente Abschnitt wird nun durch die Nebenbedingungen 3.5 und 3.6 definiert. Wird nun, damit die Nebenbedingung 3.3 bei der Evaluation der Beobachtung 3 erfüllt bleibt, der optimale Vektor der virtuellen Multiplikatoren der Inputs μ^* mit dem reziproken Wert von Ψ^A multipliziert, nimmt die Nebenbedingung 3.5 einen positiven Wert an, während die Nebenbedingung 3.6 unverändert bleibt (wieder wird die Ausdehnung der Inputs durch die reziproke Veränderung der virtuellen Multiplikatoren μ^* kompensiert). Die beiden Nebenbedingungen 3.5 und 3.6 können nicht simultan über eine Veränderung der virtuellen Multiplikatoren der Outputs ν^* bindend werden: Würden die virtuellen Multiplikatoren der Outputs ν^* verringert, damit die Nebenbedingung 3.5 bindend ist, hätte dies gleichzeitig zur Folge, dass die Nebenbedingung 3.6 negativ wird. Die einzige Möglichkeit, dass beide Nebenbedingungen simultan bindend sind, besteht darin, die Zusammensetzung des Vektors μ^* anzupassen. Die beiden Nebenbedingungen können nur dann erfüllt sein, wenn die von Beobachtung 3 relativ intensiv genutzten Inputs ein relativ kleineres Gewicht bekommen (die Umschichtung muss wie immer unter Einhaltung der Nebenbedingung 3.3 erfolgen). Beide Beobachtungen (2 und 3) bleiben effizient; die Beobachtung 3 wird allerdings dann ineffizient, wenn eine beliebige konvexe Linearkombination von anderen effizienten Beobachtungen oder eine in der Ausgangslage ineffiziente Beobachtung der Gruppe B die Beobachtung 3 dominiert. Mit zunehmender gruppenspezifischer Ineffizienz wird die typische Konstellation zu einer stabil-idealen Konstellation, bei der alle Beobachtungen des Referenzsets der Gruppe B entstammen.

Die Auswirkung der gruppenspezifischen Ineffizienz auf die ineffizienten Beobachtungen 5 und 6 lässt sich quantitativ nur anhand des in Abschnitt 2.2.3 vorgestellten dualen *Envelopment Problems* herleiten. Das Lineare Programm ist in der folgenden Gleichung explizit aufgeführt, wobei das erste Subskript die Beobachtung j , das zweite Subskript

den Input bzw. den Output kennzeichnet:

$$\begin{aligned}
\min \quad & \theta_j \\
\text{s.t.} \quad & y_{j1} - \lambda_1 y_{11} - \lambda_2 y_{21} - \lambda_3 y_{31} - \lambda_4 y_{41} - \lambda_5 y_{51} - \lambda_6 y_{61} \leq 0 \\
& \vdots \\
& y_{js} - \lambda_1 y_{1s} - \lambda_2 y_{2s} - \lambda_3 y_{3s} - \lambda_4 y_{4s} - \lambda_5 y_{5s} - \lambda_6 y_{6s} \leq 0 \\
& \theta x_{j1} \Psi^{A,B} - \lambda_1 x_{11} \Psi^A - \lambda_2 x_{21} - \lambda_3 x_{31} \Psi^A - \lambda_4 x_{41} - \lambda_5 x_{51} \Psi^A - \lambda_6 x_{61} \geq 0 \\
& \vdots \\
& \theta x_{jm} \Psi^{A,B} - \lambda_1 x_{1m} \Psi^A - \lambda_2 x_{2m} - \lambda_3 x_{3m} \Psi^A - \lambda_4 x_{4m} - \lambda_5 x_{5m} \Psi^A - \lambda_6 x_{6m} \geq 0
\end{aligned}$$

Da in dieser Konstellation die Beobachtungen 2 und 3 den relevanten Abschnitt definieren, sind nur λ_2^* und λ_3^* grösser null. Alle Terme, die $\lambda_{1,4,5,6}^*$ enthalten, können folglich eliminiert werden. Zudem wird die gruppenspezifische Ineffizienz eingeführt; x_{31} bis x_{3m} nehmen um den Faktor Ψ^A zu. Das Lineare Programm kann nun für die Beobachtung 6 (Gruppe B, ohne gruppenspezifische Ineffizienz) vereinfacht dargestellt werden:

$$\begin{aligned}
\min \quad & \theta_6 \\
\text{s.t.} \quad & y_{61} \leq + \lambda_2 y_{21} + \lambda_3 y_{31} \\
& \vdots \\
& y_{6s} \leq + \lambda_2 y_{2s} + \lambda_3 y_{3s} \\
& \theta x_{61} \geq + \lambda_2 x_{21} + \lambda_3 x_{31} \Psi^A \\
& \vdots \\
& \theta x_{6m} \geq + \lambda_2 x_{2m} + \lambda_3 x_{3m} \Psi^A
\end{aligned}$$

Es ergibt sich folgendes Bild: Da bei kleinen Änderungen das Verhältnis λ_2^* zu λ_3^* konstant bleibt, nimmt für den ersten Input der Wert der Summe von $\lambda_2^* x_{21}$ und $\lambda_3^* x_{31}$ um den Faktor $\frac{\lambda_2^* x_{21} + \lambda_3^* x_{31} \Psi^A}{\lambda_2^* x_{21} + \lambda_3^* x_{31}}$ zu.²⁰ Dasselbe gilt für die Nebenbedingungen der übrigen Inputs 2, ..., m. Da die Inputs der Beobachtung 6 im Vergleich zur Ausgangslage unverändert bleiben ($\Psi^B = 0$), muss die linke Seite des vereinfachten Linearen Programms ebenfalls um denselben Faktor zunehmen:²¹

$$\theta_j^{*\Psi} = \theta_j^* \left(\Psi^A \frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} + \frac{\lambda_2^*}{\lambda_2^* + \lambda_3^*} \right) \geq \theta_j^* \quad j = 6. \quad (3.16)$$

²⁰Die optimalen Gewichte λ^* der Linearkombination, die im *Envelopment Problem* zum minimalen θ^* führen, verschieben sich zuungunsten der ineffizienter werdenden Beobachtung. m.a.W. der Anteil des Gewichts der Beobachtung 3, $\frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*}$, nimmt mit zunehmender gruppenspezifischer Ineffizienz ab.

²¹Anhand des Strahlensatzes lässt sich zeigen, dass $\frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} = \frac{\lambda_3^* x_{3i}}{\lambda_2^* x_{2i} + \lambda_3^* x_{3i}}$ für jeden Input i erfüllt ist.

Da der Klammerausdruck nur Werte grösser gleich eins annehmen kann, wird die Effizienz der Beobachtungen der Gruppe B nie abnehmen. Im Extremfall mit $\frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} = 0$ (folglich ist $\frac{\lambda_2^*}{\lambda_2^* + \lambda_3^*} = 1$) hat die gruppenspezifische Ineffizienz keinen Einfluss auf die optimale Lösung θ_j^* . Die Gleichung 3.16 reduziert sich dann zu Gleichung 3.14. Im zweiten Extremfall ($\frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} = 1$) erhöht die gruppenspezifische Ineffizienz die optimale Lösung θ_j^* der Beobachtung 6 proportional um den Faktor Ψ^A .²² Die Gleichung 3.16 reduziert sich dann zu Gleichung 3.13. Diese beiden Fälle definieren somit die untere und obere Schranke für die Übertragung der gruppenspezifischen Ineffizienz auf die optimale Lösung.²³ Zwischen den beiden Extremen verändert sich die Lösung folglich unterproportional zu Ψ^A .

Für die Beobachtung 5 ist das folgende Lineare Programm relevant:

$$\begin{aligned}
 \min \quad & \theta_6 \\
 \text{s.t.} \quad & y_{51} \leq + \lambda_2 y_{21} + \lambda_3 y_{31} \\
 & \vdots \\
 & y_{5s} \leq + \lambda_2 y_{2s} + \lambda_3 y_{3s} \\
 & \theta x_{51} \Psi^A \geq + \lambda_2 x_{21} + \lambda_3 x_{31} \Psi^A \\
 & \vdots \\
 & \theta x_{5m} \Psi^A \geq + \lambda_2 x_{2m} + \lambda_3 x_{3m} \Psi^A
 \end{aligned}$$

Auf den linken Seiten der Nebenbedingungen sind nun die Outputs und Inputs der zu evaluierenden Beobachtung 5 aufgeführt. Die Inputs 1, ..., m werden mit der gruppenspezifischen Ineffizienz Ψ^A multipliziert. Die Summe der beiden Referenzinputs auf der rechten Seite der jeweiligen Nebenbedingung nimmt wieder um den Faktor $(\Psi^A \frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} + \frac{\lambda_2^*}{\lambda_2^* + \lambda_3^*})$ zu. Im Gegensatz zur Beobachtung 6 nimmt bei der Evaluation der Beobachtung 5 die linke Seite der Nebenbedingung der Inputs im vereinfachten *Envelopment Problem* ebenfalls um den Faktor Ψ^A zu:

$$\theta_5^{*\Psi} \Psi^A = \theta_5^* (\Psi^A \frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} + \frac{\lambda_2^*}{\lambda_2^* + \lambda_3^*}). \quad (3.17)$$

Die Division von Gleichung 3.17 durch Ψ^A ergibt:

$$\theta_j^{*\Psi} = \frac{1}{\Psi^A} \theta_j^* (\Psi^A \frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} + \frac{\lambda_2^*}{\lambda_2^* + \lambda_3^*}) \leq \theta_j^* \quad j = 5. \quad (3.18)$$

²²Dies gilt, bis die Beobachtung 6 effizient und die Beobachtung 3 ineffizient wird. Dann ist der entsprechende effiziente Abschnitt ausschliesslich durch Beobachtungen der Gruppe B definiert und somit unabhängig von gruppenspezifischer Ineffizienz.

²³Zudem wird ersichtlich, dass in den beiden Extremfällen die gruppenspezifische Ineffizienz im gleichen Ausmass auf die optimale Lösung übertragen wird wie in der nicht-idealen bzw. idealen Konstellation.

Die optimale Lösung θ_j^* nimmt folglich im Vergleich zur Ausgangslage ohne gruppenspezifische Ineffizienz einen niedrigeren Wert an. Wieder sind zwei Extremfälle zu unterscheiden. Ist $\frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} = 0$, senkt die gruppenspezifische Ineffizienz Ψ^A den Wert der optimalen Lösung θ_5^* um den reziproken Wert von Ψ^A (womit Gleichung 3.18 identisch ist mit Gleichung 3.11). Ist dagegen $\frac{\lambda_3^*}{\lambda_2^* + \lambda_3^*} = 1$, heben sich die Ψ s im Nenner und Zähler gegenseitig auf und die gruppenspezifische Ineffizienz hat keinen Einfluss auf die optimale Lösung von θ_5^* . Auch hier wird ersichtlich, dass in diesem Fall die Gleichung 3.18 der Gleichung 3.15 entspricht. Wieder stellen die beiden Extremfälle Schranken dar, in allen anderen Fällen überträgt sich die gruppenspezifische Ineffizienz unterproportional zu $\frac{1}{\Psi^A}$.

Auch wenn diese beiden Resultate nur für kleine Veränderungen der Daten, d.h. für eine relativ kleine gruppenspezifische Ineffizienz, Gültigkeit haben, lässt sich eine Aussage ableiten: Je grösser der Anteil der Gewichte $\sum \lambda_{\in A}^*$ der Referenzbeobachtungen der Gruppe A an der Summe der Gewichte $\sum \lambda^*$ ist, desto kleiner sind die Auswirkungen der gruppenspezifischen Ineffizienz auf die optimale Lösung der DEA und *vice versa*. Diese Erkenntnis lässt sich anhand der beiden Extremfälle, die ideale bzw. die nicht-ideale Konstellation, illustrieren. In der idealen Konstellation ist der Anteil der Gewichte $\sum \lambda_{\in A}^*$ einer Beobachtung der Gruppe A an der Summe der Gewichte $\sum \lambda^*$ gleich null. Folglich ist der Anteil der Gewichte $\sum \lambda_{\in B}^*$ gleich eins. Dies bedeutet, dass alle Beobachtungen des Referenzsets der Gruppe B angehören. In der nicht-idealen Konstellation ist dagegen der Anteil der Gewichte $\sum \lambda_{\in A}^*$ an der Summe der Gewichte gleich eins: Alle Beobachtungen des Referenzsets gehören zur Gruppe A.

Wird die gruppenspezifische Ineffizienz weiter erhöht, wird die effiziente Beobachtung der Gruppe A ineffizient und es tritt die oben erläuterte stabil-ideale Konstellation auf.

Folgerung 3.8 *In der typischen Konstellation wird die gruppenspezifische Ineffizienz teilweise erkannt. Die ineffizienten Beobachtungen der von der gruppenspezifischen Ineffizienz betroffenen Gruppe werden ineffizienter, während die ineffizienten Beobachtungen der anderen Gruppe effizienter werden: Die Ineffizienzunterschiede nehmen folglich zu. In welchem Ausmass die gruppenspezifische Ineffizienz erkannt wird, hängt von den Daten (bzw. von der Verteilung der optimalen Gewichte λ^*) ab und kann erst a posteriori quantifiziert werden.*

Die Resultate der mathematischen Analyse bestätigen die bisherigen Erkenntnisse aus dem Abschnitt 3.2. Die gruppenspezifische Ineffizienz kann in der nicht-idealen Konstellation nicht erkannt werden: Diese Situation tritt auf, wenn eine Gruppe eigene effiziente

Abschnitte auf der GPM definiert. Solche Abschnitte sind vollständig von den Beobachtungen der übrigen Gruppen unabhängig. Die gruppenspezifische Ineffizienz wird dagegen in der stabil- und instabil-idealen Konstellation vollständig identifiziert. Zwischen diesen beiden Extremfällen liegt die typische Konstellation: Die gruppenspezifische Ineffizienz kann teilweise erkannt werden.

3.3.4 Fazit zur statistischen Fundierung der DEA

Wird die gruppenspezifische Ineffizienz nicht oder nur teilweise erkannt, sind die Effizienzdifferenzen nicht valide. Diese Gefahr besteht, wenn sich die Rahmenbedingungen für die verschiedenen Gruppen von Beobachtungen unterscheiden und zu systematisch unterschiedlichen Input-Outputkombinationen führen. Die implizite Konsistenzbedingung (die Beobachtungen sind unabhängig und identisch in der Technologiemenge verteilt) ist dann verletzt. Somit ist nicht sichergestellt, dass die Verzerrungen im Erwartungswert für alle Gruppen von Beobachtungen gleich gross sind und die Effizienzdifferenzen sind nicht bzw. nur zufällig valide.

Die *erweiterte explizite Bedingung*, die unabhängig von der impliziten Konsistenzbedingung zu konsistenten Schätzern der Effizienz führt, lautet: Die DEA ist dann eine konsistente Methode zur Schätzung der Effizienz, wenn die Wahrscheinlichkeit einer fast vollständig effizienten Beobachtung nicht auf die Grundgesamtheit insgesamt, sondern auf *jede* Teilmenge der Grundgesamtheit, die eine Technologie (einen Abschnitt der GPM) definiert, bezogen ist. Nur so ist gewährleistet, dass die Schätzer der DEA auch unter den abgeschwächten Annahmen (die Beobachtungen sind nicht unabhängig und identisch in der Technologiemenge verteilt) konsistent sind. Eine Argumentation könnte darauf abzielen, dass mit zunehmender Stichprobengrösse und besonders im Extremfall einer unendlich grossen Stichprobe sichergestellt ist, dass jeder Abschnitt der GPM durch vollständig effiziente Beobachtungen definiert wird. Dieses Argument trifft allgemein aber nur unter der erweiterten expliziten Konsistenzbedingung zu, wonach die zu Grunde liegenden wahren Dichtefunktionen der Effizienz für *alle* Gruppen von Beobachtungen (d.h. der Gruppe A und B) am oberen Ende des Definitionsbereiches²⁴ strikt positive Werte annehmen – eine Annahme, die – wie die ursprüngliche explizite Konsistenzbedingung – nicht überprüft werden kann.

Die Bedingung für (asymptotisch) valide Effizienzdifferenzen lautet: Die geschätzte Ef-

²⁴Das obere Ende des Definitionsbereichs ist durch vollständige Effizienz definiert.

fizienz ist im Erwartungswert gleich stark verzerrt, wenn die Wahrscheinlichkeit einer fast vollständig effizienten Beobachtung für *jede* Teilstichprobe, die eine Technologie definiert, gleich gross ist.

Die Konsistenz der Schätzer der Effizienz und die Validität der geschätzten Effizienzdifferenzen der DEA sind an Annahmen über die Effizienzverteilungen geknüpft; Annahmen, die im Rahmen einer empirischen Forschung nicht überprüft werden können. Es besteht folglich die latente Gefahr unterschiedlich stark verzerrt geschätzter Effizienz und nicht valider Effizienzdifferenzen. Wie soll in der empirischen Forschung mit diesen Situationen umgegangen werden? Diese Frage ist Gegenstand des nächsten Abschnittes.

3.4 Konsistenzprobleme der DEA in der empirischen Forschung

3.4.1 Identifikation nicht-idealer Konstellationen

Die Frage stellt sich, ob im Rahmen der empirischen Forschung nicht-ideale Konstellationen identifiziert werden sollen. Wird darauf verzichtet, besteht eine erhöhte Wahrscheinlichkeit, dass Hypothesen zu unrecht abgewiesen oder zu unrecht nicht abgewiesen werden. Solche potentiellen Fehler erster oder zweiter Art stellen die Resultate einer Effizienzmessung im Rahmen einer empirischen Forschung in Frage. Die Kosten solcher Fehler dürften gerade in volkswirtschaftlichen Anwendungen den Zusatzaufwand rechtfertigen.

Wie sollen aber nicht-ideale Konstellationen identifiziert werden, und wie kann die Gesamtheit aller effizienten Abschnitte der GPM beurteilt werden? Grundsätzlich bestehen zwei Möglichkeiten, um gruppenspezifisch unabhängige effiziente Abschnitte (nicht-ideale Konstellationen) zu identifizieren. Die erste besteht darin, anhand statistischer Methoden nach systematischen Zusammenhängen zwischen Variablen, die die Rahmenbedingungen erfassen, und Inputs und Outputs zu suchen. Ein eventueller Zusammenhang kann z.B. anhand einer Korrelationsanalyse festgestellt werden: Eine starke Korrelation zwischen den Input-Outputtransformationen und den die Rahmenbedingungen erfassenden Variablen ist ein Indiz für die Existenz von nicht-idealen Konstellationen.

Die zweite Variante besteht darin, direkt zu untersuchen, welche Beobachtungen die effizienten Abschnitte definieren.²⁵ Eine Auswertung der Zusammensetzung der GPM,

²⁵Die dazu notwendigen Informationen sind Teil der Lösung der Linearen Programme, welche aber –

welche aus einer Vielzahl einzeln zu bewertender Abschnitte bestehen kann, ist eine aufwendige Angelegenheit. Zudem stellt sich die Frage, wie die GPM insgesamt beurteilt werden soll, falls sich diese aus idealen, nicht-idealen und typischen Konstellationen zusammensetzt; in Abschnitt 3.3 wurde immer nur jeweils eine Konstellation analysiert. Ein Beispiel soll zeigen, wie sich eine Änderung der Inputvektoren auf die *durchschnittliche* Effizienz der beiden Gruppen auswirkt, falls verschiedene Konstellationen auftreten. Der Anteil der Beobachtungen der Gruppe A in stabil-idealen Konstellationen sei α , die übrigen Beobachtungen seien in der instabil-idealen oder nicht-idealen Konstellation vertreten. Nach einer Änderung der Inputvektoren der Gruppe A um den Faktor Ψ^A wird die durchschnittliche Effizienz $\bar{z}_A^{*\Psi^A}$ der Gruppe A neu

$$\bar{z}_A^{*\Psi^A} = \frac{\alpha}{\Psi^A} * \bar{z}_A^* + (1 - \alpha) * \bar{z}_A^* \quad (3.19)$$

betragen (vgl. Folgerungen 3.5 und 3.7). Je nach dem wie stark die einzelnen Konstellationen vertreten sind, bleibt die durchschnittliche Effizienz $\bar{z}_A^{*\Psi^A}$ unverändert ($\alpha = 0$), verändert sich partiell ($0 < \alpha < 1$) oder vollständig, d.h. um den Faktor $\frac{1}{\Psi^A}$ ($\alpha = 1$). In diesem Beispiel sei zudem der Anteil der Beobachtungen der Gruppe B in instabil-idealen β , der Rest sei in stabil-idealen bzw. nicht-idealen Konstellationen vertreten. Die durchschnittliche Effizienz $\bar{z}_B^{*\Psi^A}$ der Gruppe B entspricht nach der Veränderung der Inputvektoren der Gruppe A

$$\bar{z}_B^{*\Psi^A} = \beta * \Psi^A * \bar{z}_B^* + (1 - \beta) * \bar{z}_B^*. \quad (3.20)$$

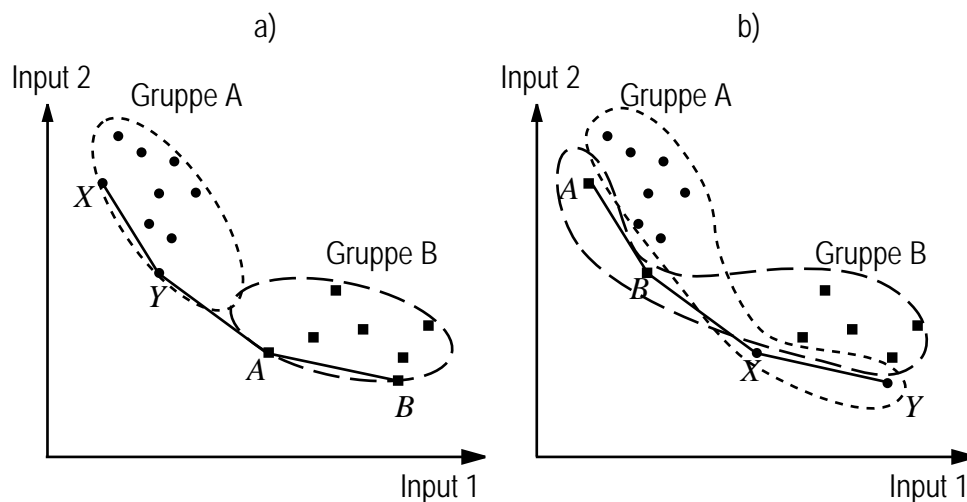
Die durchschnittliche Effizienz der Gruppe B bleibt für $\beta = 0$ unverändert, verändert sich partiell ($0 < \beta < 1$) oder vollständig, d.h. um den Faktor Ψ^A falls $\beta = 1$ gilt (vgl. Folgerungen 3.6 und 3.7). Die Änderung der Inputvektoren der Gruppe A wirkt sich in Abhängigkeit der Verteilung der stabil-idealen, instabil-idealen, nicht-idealen und typischen Konstellation unterschiedlich stark auf die durchschnittliche Effizienz der beiden Gruppen aus.

Der zweiten Variante, der Analyse der Zusammensetzung der GPM, sollte der Vorzug gegeben werden, da anhand der Korrelationskoeffizienten lediglich zentrale Zusammenhänge erkannt werden können. Einige wenige Ausreisser wirken sich nur marginal auf den Wert des Korrelationskoeffizienten aus. In der DEA definieren aber einzelne extreme Beobachtungen die effizienten Abschnitte. Der Korrelationskoeffizient, der zentrale

im Gegensatz zur ersten Variante – erst nach erfolgter Effizienzmessung vorliegen.

Zusammenhänge der Daten erfasst, eignet sich nicht zur Beschreibung einer Teilmenge der Daten, die *systematisch* unterschiedliche Eigenschaften aufweist. Die effizienten Beobachtungen sind aber eine in diesem Sinne extreme Teilmenge aller Beobachtungen. Eine geringe Korrelation zwischen dem Inputverhältnis und der Gruppenzugehörigkeit garantiert daher noch nicht eine ideale Konstellation und *vice versa*. In der Abbildung 3.6a) ist die nicht-ideale Konstellation dargestellt; das Inputverhältnis der Beobachtungen ist mit der Gruppenzugehörigkeit korreliert. In der Abbildung 3.6b) ist eine ideale Konstellation dargestellt. Wieder sind die Inputverhältnisse mit der Gruppenzugehörigkeit korreliert – im Gegensatz zur Abbildung 3.6a) definieren nun aber Beobachtungen der Gruppe A das Referenzset der ineffizienten Beobachtungen der Gruppe B und *vice versa*. Auch wenn der Korrelationskoeffizient kleiner ausfallen müsste als in der nicht-idealen Konstellation, lässt er i.d.R. keinen eindeutigen Rückschluss auf die Existenz nicht-idealer Konstellationen zu.

Abbildung 3.6: Mit der Gruppenzugehörigkeit korrelierte Inputverhältnisse



3.4.2 Eine Kennzahl zur Bewertung der Daten: Das Validitätsmass der Effizienzdifferenzen

Wie erwähnt wurde, ist es im Rahmen einer empirischen Effizienzmessung nicht möglich, die Konsistenzbedingungen zu überprüfen und die Verzerrung der geschätzten Effizienz zu ermitteln. Für die Hypothesentests bezüglich der Effizienzmittelwerte von Gruppen von Beobachtungen ist aber nicht das Niveau der Effizienz, das durch die Verzerrung beeinflusst wird, sondern lediglich die Effizienzdifferenz relevant. Die Validität der Effizienzdifferenz lässt sich im Rahmen einer empirischen DEA mit der im Folgenden entwickelten Kennzahl ermitteln.

Im Rahmen einer Sensitivitätsanalyse lässt sich für jede Gruppe von Beobachtungen die relative Veränderung der durchschnittlichen Effizienz \bar{z}^* bezüglich einer Veränderung der gruppenspezifischen Ineffizienz Ψ^A bestimmen. Dabei werden die Inputs der zu evaluierenden Untergruppe A mit dem Faktor Ψ^A multipliziert, die DEA wird wiederholt und $\Delta_{\Psi^A}^A$, der Anteil der effektiv realisierten relativen Veränderung der durchschnittlichen Effizienz der Gruppe A an der maximal möglichen relativen Veränderung der Effizienz der Gruppe A (infolge der Veränderung der Inputvektoren der Gruppe A), wird gemäss Gleichung 3.21 berechnet:

$$\Delta_{\Psi^A}^A = \frac{\frac{\bar{z}_A^* - \bar{z}_A^{*\Psi^A}}{\bar{z}_A^*}}{1 - \frac{1}{\Psi^A}} \quad \forall j \in A. \quad (3.21)$$

In Gleichung 3.21 ist \bar{z}_A^* die durchschnittliche Effizienz der Beobachtungen der Gruppe A vor der Veränderung der Inputvektoren. Dagegen entspricht $\bar{z}_A^{*\Psi^A}$ der durchschnittlichen Effizienz nach der Veränderung der Daten. Der Zähler in Gleichung 3.21 entspricht somit der relativen Veränderung der durchschnittlichen Effizienz der Beobachtungen der Gruppe A infolge der Veränderung der Inputvektoren um den Faktor Ψ^A . Der Nenner ist gleich der wahren relativen Veränderung der durchschnittlichen Effizienz infolge der Veränderung der Inputvektoren, d.h. die Veränderung, die in der stabil-idealen Konstellation realisiert würde.²⁶ Durch diese Normierung ist gewährleistet, dass $\Delta_{\Psi^A}^A$ auf Werte zwischen null und eins beschränkt ist. Die Kennzahl $\Delta_{\Psi^A}^A$ gibt den prozentualen Anteil einer Veränderung der Inputvektoren an, der sich auf die durchschnittliche Effizienz \bar{z}_A^* der Gruppe A überträgt. Ein Wert von eins wird realisiert, wenn alle Beobachtungen

²⁶Da eine proportionale Veränderung aller Inputs zu einer entsprechenden reziproken Veränderung der optimalen Lösung führt, wird im Nenner der reziproke Wert von Ψ^A verwendet (vgl. Folgerung 3.5).

der Gruppe A in stabil-idealen Konstellationen auftreten, während ein Wert von null auf ausschliesslich nicht-ideale Konstellationen hinweist.²⁷ In Abbildung 3.7a) ist eine stabil-ideale Konstellation dargestellt: Die beiden Beobachtungen X und Y der Gruppe B definieren das Referenzset der Beobachtungen der Gruppe A. Die ursprüngliche Situation der Gruppe A ist grau dargestellt. Die Veränderung der Inputs führt zu einer reziproken Abnahme der durchschnittlichen Effizienz der Beobachtungen der Gruppe A und $\Delta_{\Psi^A}^A$ nimmt folglich einen Wert von eins an.

Die Folgerung 3.6 besagt, dass in der instabil-idealen Konstellation die Effizienzwerte der Gruppe A unverändert bleiben, eine Veränderung der Inputvektoren der Gruppe A aber die Effizienzwerte der Gruppe B beeinflusst und somit für die Effizienzdifferenz zwischen den beiden Gruppen relevant ist. Da davon ausgegangen werden muss, dass in einer empirischen Anwendung alle Konstellationen – d.h. auch die instabil-ideale – simultan auftreten, muss die Veränderung der durchschnittlichen Effizienz der Gruppe B infolge einer Veränderung der Inputvektoren der Gruppe A berücksichtigt werden. Die instabil-ideale Konstellation ist in Abbildung 3.7b) dargestellt; die beiden Beobachtungen C und D der Gruppe A definieren das Referenzset für die ineffizienten Beobachtungen der Gruppe B. Eine Veränderung der Inputs der Gruppe A führt nicht zu einer Veränderung der geschätzten Effizienz der Beobachtungen C und D : In Gleichung 3.21 ist $\bar{z}_A^* = \bar{z}_A^{*\Psi^A}$ und folglich nimmt $\Delta_{\Psi^A}^A$ einen Wert von null an. Die Veränderung der Inputvektoren der Gruppe A hat in Abbildung 3.7b) aber Konsequenzen für die Beobachtungen der Gruppe B: Die Effizienz der Beobachtungen der Gruppe B nimmt proportional zu Ψ^A zu (vgl. Folgerung 3.6). Der Anteil der effektiv realisierten relativen Veränderung der durchschnittlichen Effizienz der Gruppe B an der maximal möglichen relativen Veränderung der Effizienz der Gruppe B infolge einer Veränderung der Inputvektoren der Gruppe A entspricht

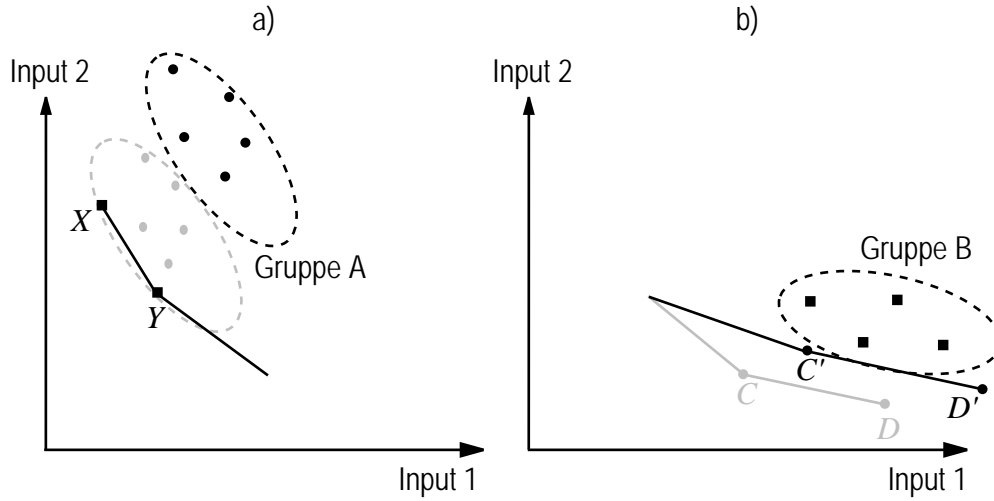
$$\Delta_{\Psi^A}^B = -\frac{\frac{\bar{z}_B^* - \bar{z}_B^{*\Psi^A}}{\bar{z}_B^*}}{1 - \Psi^A} \quad \forall j \in B. \quad (3.22)$$

²⁷Dass der Wertebereich von $\Delta_{\Psi^A}^A$ auf einem Intervall zwischen null und eins beschränkt ist wird ersichtlich, wenn $\bar{z}_A^{*\Psi^A}$ in Gleichung 3.21 durch die rechte Seite von Gleichung 3.19 ersetzt wird:

$$\Delta_{\Psi^A}^A = \frac{\bar{z}_A^* - (\frac{\alpha}{\Psi^A} * \bar{z}_A^* + (1 - \alpha) * \bar{z}_A^*)}{\bar{z}_A^*} / (1 - \frac{1}{\Psi^A}).$$

Kürzen ergibt den folgenden Ausdruck: $\Delta_{\Psi^A}^A = (\alpha - \frac{\alpha}{\Psi^A}) / (1 - \frac{1}{\Psi^A})$, der für $\alpha = 0$ den Wert null und für $\alpha = 1$ den Wert eins annimmt.

Abbildung 3.7: Die beiden Komponenten des Validitätsmasses



$\Delta_{\Psi^A}^B$ ist dabei auf Werte zwischen minus eins und null beschränkt.²⁸ Ein Wert von minus eins wird realisiert, wenn die Beobachtungen der Gruppe B ausschliesslich in instabil-idealen Konstellationen auftreten [wie in der Abbildung 3.7b)], ein Wert von null dagegen ist ein Indiz dafür, dass ausschliesslich nicht-ideale Konstellationen vorliegen.²⁹

Ist die GPM vorwiegend durch stabil-ideale Konstellationen gekennzeichnet, wird $\Delta_{\Psi^A}^A$ im Verhältnis zu $\Delta_{\Psi^A}^B$ einen grossen Wert annehmen. Überwiegt dagegen die instabil-ideale

²⁸Dass der Wertebereich von $\Delta_{\Psi^A}^B$ auf einem Intervall zwischen minus eins und null beschränkt ist wird ersichtlich, wenn $\bar{z}_B^{*\Psi^A}$ in Gleichung 3.22 durch die rechte Seite von Gleichung 3.20 ersetzt wird:

$$\Delta_{\Psi^A}^B = -\frac{\bar{z}_B^* - (\beta * \Psi^A * \bar{z}_B^* + (1 - \beta) * \bar{z}_B^*)}{\bar{z}_B^*} / (1 - \Psi^A).$$

Kürzen ergibt den folgenden Ausdruck: $\Delta_{\Psi^A}^B = -\beta(1 - \Psi^A)/(1 - \Psi^A)$, der für $\beta = 0$ den Wert null und für $\beta = 1$ den Wert minus eins annimmt.

²⁹Der Zähler in Gleichung 3.22 erfasst in Analogie zu Gleichung 3.21 die durchschnittliche Effizienz vor und nach der Änderung der Inputvektoren der Gruppe A, bezieht sich allerdings auf die durchschnittliche Effizienz der Gruppe B. Die beiden Nenner, die die maximal möglichen Veränderungen der durchschnittlichen Effizienz ausweisen, unterscheiden sich dagegen: In Folgerung 3.5 wurde festgehalten, dass eine Veränderung der Inputvektoren der Gruppe A einen reziproken Effekt auf die Effizienz hat. Daher ist im Nenner der Gleichung 3.21 Ψ^A reziprok berücksichtigt. Eine Aussage der Folgerung 3.6 ist, dass eine Veränderung der Inputvektoren einen proportionalen Effekt auf die Effizienz hat – daher wird Ψ^A in Gleichung 3.22 entsprechend berücksichtigt.

Konstellation, wird $\Delta_{\Psi A}^B$ einen Wert nahe bei minus eins annehmen, während für $\Delta_{\Psi A}^A$ ein Wert um null realisiert wird. Ist die GPM mehrheitlich durch nicht-idealen Konstellation gekennzeichnet, nimmt sowohl $\Delta_{\Psi A}^A$ wie auch $\Delta_{\Psi A}^B$ einen Wert um null an.

Relevant ist, dass eine Veränderung der Inputvektoren der Gruppe A einen Einfluss auf die geschätzte Effizienz der Gruppe A und B haben kann. Damit Aussagen zur Validität der ermittelten Effizienzdifferenz gemacht werden können, ist es daher notwendig, die Veränderung der durchschnittlichen Effizienz *beider* Gruppen zu berücksichtigen. Im Folgenden wird die Validität der Effizienzdifferenz zwischen den Gruppen A und B durch das *Validitätsmass der Effizienzdifferenz* des Gruppenmerkmals A-B³⁰ erfasst, das der Summe der absoluten Werte von $\Delta_{\Psi A}^A$ und $\Delta_{\Psi A}^B$ entspricht:

$$\Delta_{\Psi A}^{AB} = \Delta_{\Psi A}^A + |\Delta_{\Psi A}^B|. \quad (3.23)$$

Das Validitätsmass $\Delta_{\Psi A}^{AB}$ ist theoretisch auf Werte zwischen null und zwei beschränkt. Nimmt das Validitätsmass einen Wert nahe bei null an, überwiegt die nicht-ideale Konstellation und es besteht die Gefahr verzerrter Effizienzdifferenzen. Eine Veränderung der Inputvektoren einer Gruppe hat dann keinen Einfluss auf die ermittelte Effizienzdifferenz. Bei Werten von rund eins überwiegt die stabil-ideale und/oder die instabil-ideale Konstellation, die Effizienzdifferenzen sind daher unverzerrt und somit valide. In dieser Situation führt eine Veränderung der Inputvektoren einer Gruppe zu einer entsprechenden Veränderung der Effizienzdifferenz. Werte zwischen eins und zwei werden ausschliesslich in ungewöhnlichen Konstellationen realisiert. Eine solche Konstellation ist in der Abbildung 3.6b) dargestellt: Wenige Beobachtungen der Gruppe A definieren das Referenzset für die überwiegende Mehrheit der Beobachtungen der Gruppe B und *vice versa*. Folglich ist $\Delta_{\Psi A}^A$ nur unwesentlich kleiner eins während $\Delta_{\Psi A}^B$ nur unwesentlich grösser minus eins ist. I.d.R. ist aber $\Delta_{\Psi A}^{AB}$ nach oben auf Werte um rund eins beschränkt, dies wird sich auch im folgenden Kapitel bei den *Monte-Carlo*-Simulationen zeigen.

Das Validitätsmass wird für jedes Gruppenmerkmal berechnet. Überträgt sich die Veränderung der Inputvektoren einer Gruppe nicht auf die Differenz der durchschnittlichen Effizienz ($\Delta_{\Psi A}^{AB} \approx 0$), sind die Effizienzdifferenzen nicht valide und es besteht die Gefahr, dass mögliche Effizienzunterschiede nicht erkannt werden. Die statistische Auswertung der Effizienz bezüglich dieses Gruppenmerkmals liefert dann keine Informationen über gruppenspezifische Ineffizienz bzw. Effizienzunterschiede. Die Konsequenz ist, dass

³⁰Als Gruppenmerkmal wird eine Variable verstanden, die eine Unterteilung der Stichprobe in zwei Teilstichproben ermöglicht. Beobachtungen sind z.B. öffentlich-rechtlich oder privatrechtlich organisiert.

Effizienzunterschiede unterschätzt werden und die Wahrscheinlichkeit zunimmt, dass die Nullhypothese zu unrecht nicht verworfen wird: Es besteht eine erhöhte Gefahr von Fehlern zweiter Art.³¹

Folgerung 3.9 *Im Rahmen einer Sensitivitätsanalyse kann überprüft werden, ob die durchschnittliche Effizienz auf eine Veränderung der gruppenspezifischen Ineffizienz reagiert. Nimmt das Validitätsmass Δ ³² in Bezug auf ein Gruppenmerkmal einen kleinen Wert an, besteht die Gefahr, dass Effizienzunterschiede unterschätzt werden.*

Es stellt sich die Frage, in welcher Verbindung das Validitätsmass Δ mit unerkannten Effizienzunterschieden steht und wie sich diese auf die Ebene der Hypothesentests übertragen. Die bisherigen Erkenntnisse legen die folgenden Hypothesen nahe:

Hypothese 1 *Nimmt das Validitätsmass Δ im Rahmen einer DEA in Bezug auf ein Gruppenmerkmal einen Wert um eins an, sind die erwarteten Effizienzdifferenzen zwischen den Gruppen valide, d.h. unverzerrt.*

Hypothese 2 *Bestehen Effizienzunterschiede zwischen zwei zu evaluierenden Gruppen, ist bei kleineren Werten von Δ mit einer erhöhten Wahrscheinlichkeit von Fehlern zweiter Art (Nichterkennen tatsächlicher Effizienzunterschiede) zu rechnen.*

Diese Hypothesen werden im folgenden Kapitel 4 anhand einer *Monte-Carlo*-Simulation überprüft.

³¹Die Frage nach den kritischen Werten von $\Delta_{\Psi A}^{AB}$ bleibt vorläufig ungeklärt und ist Gegenstand des nächsten Kapitels.

³²Im Folgenden wird das gemäss Gleichung 3.23 berechnete Validitätsmass vereinfacht als Δ bezeichnet, wobei ein Superskript als Indikator für das Gruppenmerkmal dient.

Kapitel 4

Monte-Carlo-Simulation der Effizienzanalyse

Das Ziel dieses Kapitels ist, die Hypothesen des dritten Kapitels anhand einer *Monte-Carlo*-Simulation zu überprüfen. Bei der *Monte-Carlo*-Simulation geht es darum, die Validität einer Methode zu überprüfen. Ausgehend von einem „wahren“ Modell werden Stichproben finiter Grösse generiert und anschliessend die Parameter des unterstellten Modells geschätzt. Der Vergleich zwischen den „wahren“ und den geschätzten Parametern lässt einen Rückschluss auf die Eigenschaften der zu überprüfenden Methode zu.

In dieser Simulation wird die Beziehung zwischen den Verzerrungen und dem Validitätsmass Δ analysiert. Es wird einerseits untersucht, ob das Ausmass der Verzerrungen bei der Schätzung der Effizienz mit dem Validitätsmass Δ in Verbindung steht. Andererseits interessiert, unter welchen Umständen die geschätzten Effizienzdifferenzen unverzerrt, d.h. valide sind. Valide sind Effizienzdifferenzen, wenn die Verzerrungen der geschätzten Effizienz bei den zu analysierenden Gruppen im Erwartungswert gleich gross sind. Zudem wird untersucht, ob und wie sich unterschiedlich starke Verzerrungen, d.h. nicht valide Effizienzdifferenzen, auf die Wahrscheinlichkeit der Fehler erster bzw. zweiter Art auswirken.

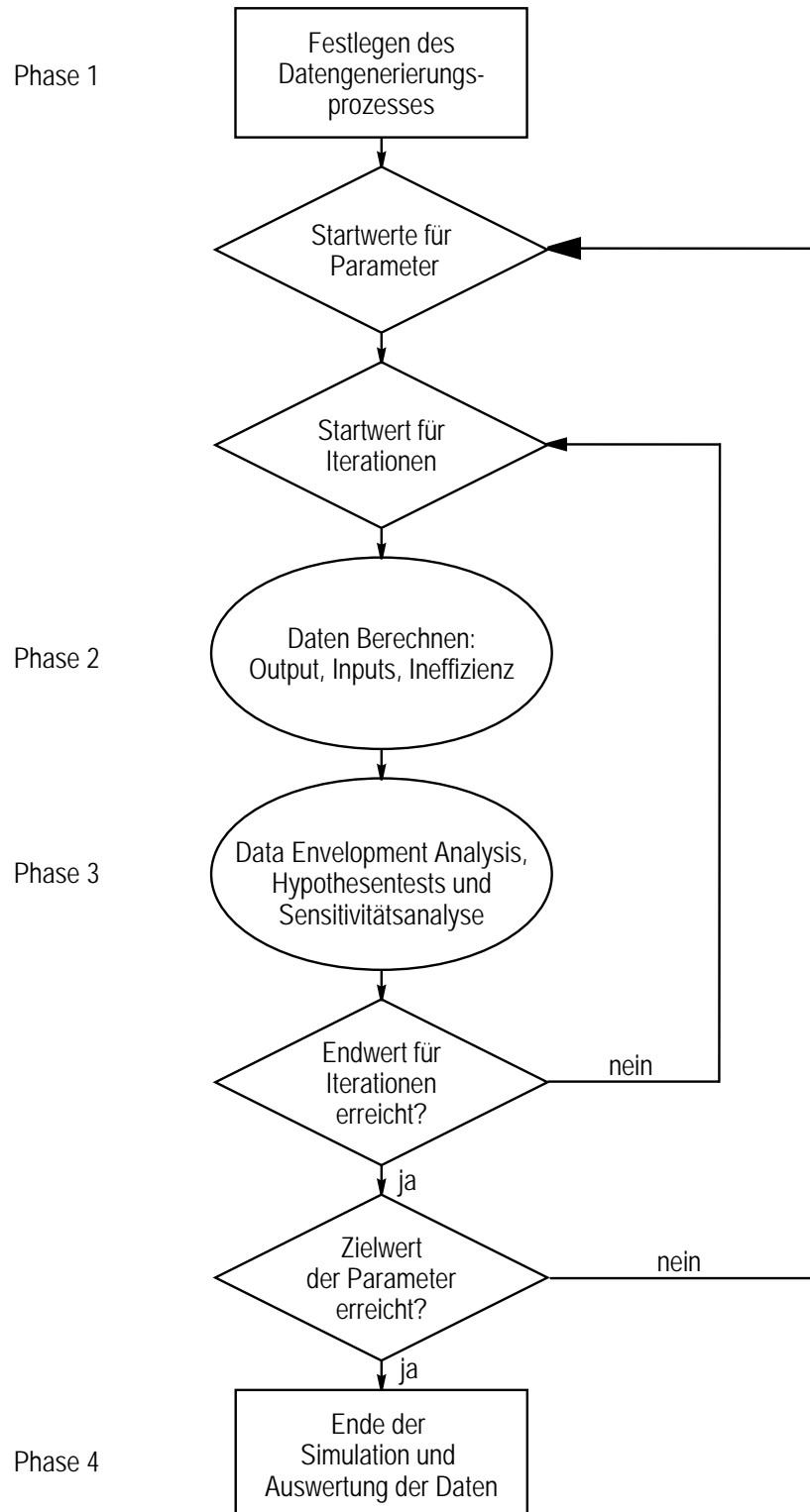
4.1 Der Ablauf der Simulation

Monte-Carlo-Simulationen lassen sich in vier Phasen gliedern. Abbildung 4.1 stellt den Ablauf einer *Monte-Carlo*-Simulation dar. In der ersten Phase wird der Datengenerie-

rungsprozess definiert (vgl. Abschnitt 4.2). Dieser Datengenerierungsprozess umfasst die Definition des „wahren“ Modells. Es wird festgelegt, in welcher Beziehung die Variablen zueinander stehen und wie sie in der Grundgesamtheit verteilt sind, d.h. den Parametern werden konkrete Werte zugeteilt. In der zweiten Phase werden Zufallszahlen gezogen und die Daten unter Berücksichtigung des „wahren“ Modells berechnet, m.a.W. es werden Stichproben generiert (vgl. Abschnitt 4.3). In der dritten Phase werden die DEA, die Hypothesentests und die Analyse der Validität durchgeführt (vgl. Abschnitt 4.4). Die zweite und dritte Phase werden solange wiederholt, bis die vorgegebene Anzahl Iterationen erreicht worden ist. Jede Iteration entspricht dabei einer vollständigen DEA mit Hypothesentests (und Analyse der Validität), wie sie in der empirischen Forschung durchgeführt werden. Die Anzahl der durchzuführenden Iterationen richtet sich nach dem Ziel der Simulation. Werden die Kosten der Simulation vernachlässigt, gilt grundsätzlich: Je mehr Iterationen desto besser. Denn aus jeder Iteration resultiert ein geschätzter Parameterwert. Werden 100 Iterationen durchgeführt, resultieren folglich 100 geschätzte Parameterwerte, die zusammen eine Verteilung (die sog. *Sampling Distribution*) des geschätzten Parameters ergeben. Mit jeder zusätzlichen Iteration wird zusätzliche Information über diese Parameterverteilung bereitgestellt. Gerade wenn die Schwänze der *Sampling Distribution* im Zentrum des Interesses stehen, muss eine sehr grosse Anzahl Iterationen durchgeführt werden, da solche extremen Werte selten realisiert werden. Eine relativ kleine Anzahl Iterationen sollte dagegen ausreichen, wenn der Mittelwert der Parameterverteilung interessiert. In dieser Simulation werden jeweils 500 Iterationen durchgeführt, womit gewährleistet sein sollte, dass auch extrem seltene Parameterwerte realisiert werden.

Nach 500 Iterationen wird der Wert *eines* Parameters um eine Einheit verändert und anschliessend wird die zweite und dritte Phase mit der neuen Parameterkombination wiederholt. Nach weiteren 500 Iterationen wird wieder der Wert *eines* Parameters um eine Einheit verändert. Insgesamt werden 270 verschiedene Parameterkombinationen simuliert. Somit ist gewährleistet, dass die Resultate der Simulation nicht nur für eine, sondern für verschiedenste Parameterkombinationen Gültigkeit haben. Nach der Simulation aller Parameterkombinationen werden in der vierten Phase die geschätzten Parameter mit den „wahren“ Parametern verglichen, die Verzerrungen berechnet und die Resultate der Hypothesentests ausgewertet (vgl. Abschnitt 4.5). Die ersten drei Phasen der Simulation werden mit MATLAB (Release 6.1) und die vierte mit SPSS (Release 10) durchgeführt. Die Syntax für die ersten drei Phasen ist im Anhang A aufgeführt.¹

¹Besonderen Dank geht an die Herren Professoren Philippe Vanden Eeckaut und Léopold Simar (Insti-

Abbildung 4.1: Flussdiagramm der *Monte-Carlo*-Simulation

4.2 Phase 1: Der Datengenerierungsprozess

Dieser Abschnitt befasst sich mit dem Datengenerierungsprozess. Es muss ein Prozess definiert werden, der die Generierung von künstlichen Input-Outputdaten in der zweiten Phase der Simulation erlaubt. Dazu muss explizit eine Transformationsfunktion festgelegt und es müssen Annahmen über die Verteilung der Effizienz getroffen werden. Zudem wird definiert, welchen Grössen zufällige Werte zugewiesen und welche berechnet werden:

1. Die allgemeine Formulierung der Technologiemenge „ Y kann mit X produziert werden“ wird durch eine konkrete Transformationsfunktion $F(Y, X) \leq 0$ ersetzt. In dieser Simulation findet eine *Cobb-Douglas*-Transformationsfunktion mit zwei Inputs und konstanten Skalenerträgen Anwendung:²

$$\mathcal{T} = \{(Y, X) | F(Y, X) \leq 0\}, \quad F(Y, X) = y - x_1^{0.5} x_2^{0.5}. \quad (4.1)$$

Die DEA findet in der empirischen Wirtschaftsforschung gerade dann Verwendung, wenn mehrere Inputs und Outputs in den Transformationsprozess eingehen. Daher wäre es wünschenswert, auch in der Simulation eine Transformationsfunktion mit mehreren Inputs und Outputs zu verwenden. Allerdings kann das gestellte Ziel dieser Arbeit mit dieser einfachen Transformationsfunktion erreicht werden.³ Wie in Abschnitt 3.1 erläutert wurde, sind die DEA-Schätzer der Effizienz verzerrt aber unter Umständen konsistent. Kneip et al. (1998) haben zudem gezeigt, dass die geschätzte Effizienzverteilung $\hat{\theta}_{DEA}$ sich der wahren Verteilung θ_{DEA} mit der Konvergenzrate

$$O = n^{-\frac{2}{s+m+1}} \quad (4.2)$$

annähert (wobei s die Anzahl Outputs, m die Anzahl Inputs und n die Anzahl Beobachtungen repräsentiert). Mit zunehmender Stichprobengrösse bzw. mit abnehmender Anzahl Dimensionen des Input-Outputraumes nähert sich die geschätzte Effizienzverteilung der „wahren“ Effizienzverteilung an. Je kleiner die Anzahl Inputs

tute de Statistique, Université Catholique de Louvain), die ihren bewährten MATLAB Code zur Lösung der DEA für meine MATLAB *Monte-Carlo*-Syntax zur Verfügung stellten.

²Die *Cobb-Douglas*-Transformationsfunktion wurde gewählt, da sie bekannt und einfach zu berechnen ist. Andere Transformationsfunktionen sind auch denkbar und es ist nicht wahrscheinlich, dass die Wahl der Transformationsfunktion einen Einfluss auf die Resultate hat.

³Zudem wird im Abschnitt 4.6 der Versuch unternommen, die Resultate auf den allgemeinen Fall zu übertragen.

und Outputs ist, desto kleiner sind bei gegebener Stichprobengrösse die erwarteten Verzerrungen.

2. Der Output y_j und der Inputmix $\frac{x_{j1}}{x_{j1}+x_{j2}}$, d.h. der Anteil des ersten Inputs an der Summe der Inputs, stellen die beiden Variablen der Produktionsfunktion dar, die mit Zufallszahlen generiert werden. Der Output und der Inputmix sind im wahren Modell normalverteilte Zufallsvariablen mit *a priori* festgelegten Erwartungswerten μ_y und μ_x und Standardabweichungen σ_y und σ_x :

$$\begin{aligned} y_j &\sim N(\mu_y, \sigma_y) \\ \frac{x_{j1}}{x_{j1} + x_{j2}} &\sim N(\mu_x, \sigma_x). \end{aligned}$$

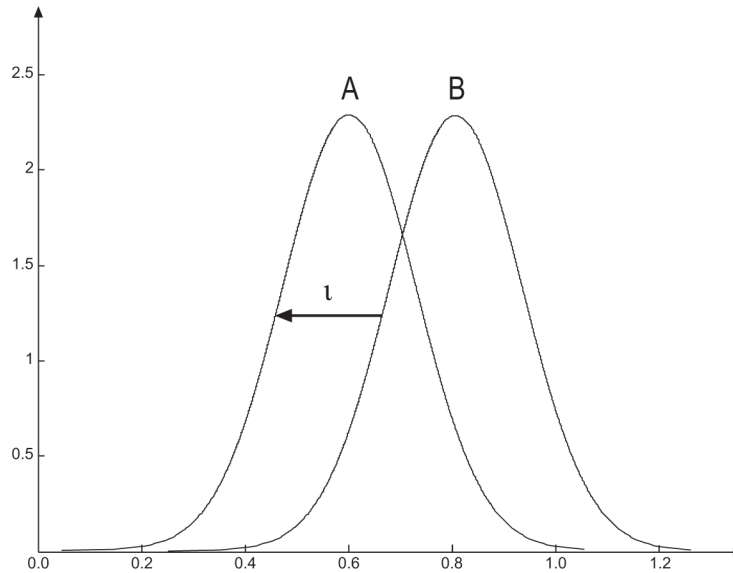
Diese beiden Informationen lassen in der zweiten Phase die Berechnung beider Inputmengen (x_1 und x_2) zu. Diese Vorgehensweise mag auf den ersten Blick etwas umständlich erscheinen. Schliesslich wäre es grundsätzlich auch möglich, die beiden Inputs direkt durch das Ziehen von Zufallszahlen zu bestimmen und dann den Output gemäss Gleichung 4.1 zu berechnen. Da aber in dieser Simulation gerade der Inputmix im Zentrum des Interesses steht und bewusst variiert werden soll, wird auf diese naheliegendere Vorgehensweise verzichtet.⁴

3. Da alle dermassen generierten „Beobachtungen“ auf der GPM liegen, wird den „Beobachtungen“ künstlich Ineffizienz hinzugefügt. Die Effizienz i ist normalverteilt mit *a priori* bekanntem Erwartungswert μ_i und Standardabweichung σ_i . Da die Effizienz i und die überschüssigen Inputs in einem reziproken Verhältnis stehen, werden in der zweiten Phase die Inputmengen jeder „Beobachtung“ j durch den Effizienzfaktor i_j dividiert:

$$\begin{aligned} i_j &\sim N(\mu_i, \sigma_i) \\ \hat{x}_{j1} &= x_{j1} i_j^{-1} \\ \hat{x}_{j2} &= x_{j2} i_j^{-1}. \end{aligned}$$

Auch diese Vorgehensweise mag etwas umständlich wirken, könnte die reduzierte Effizienz doch direkt vermindernd auf den Output wirken. Da aber die gruppenspezifische Ineffizienz Gegenstand der Untersuchung ist und folglich für jede Gruppe

⁴Zudem ergäben sich unter Umständen bei dieser Vorgehensweise systematische Unterschiede in den Outputniveaus der zu untersuchenden Gruppen. Im Falle der DEA mit variablen Skalenerträgen würde dies eine Verzerrung der Simulationsergebnisse nach sich ziehen.

Abbildung 4.2: Illustration des Verschiebungsparameters ι 

unterschiedlich ausfallen soll, würden wieder unerwünschte Effekte resultieren: Weniger effiziente Gruppen von Beobachtungen würden dann bei der direkten Methode ein systematisch geringeres Outputniveau aufweisen. Mit der hier verwendeten Vorgehensweise wird sichergestellt, dass die beiden zu untersuchenden Gruppen im Rahmen der definierten Verteilung des Outputs im Erwartungswert identische Outputmengen produzieren. Die Normalverteilung bietet sich als Effizienzverteilung an, da somit in der dritten Phase der Simulation der T-Test zur Überprüfung der Hypothesen verwendet werden kann.

Schliesslich müssen noch die in der Simulation schrittweise zu verändernden Parameter definiert werden. Ein erster Parameter ist die gruppenspezifische Ineffizienz ι , welche die Grundlage für die Hypothesentests darstellt, d.h. die Beantwortung der Frage erlaubt, ob die gruppenspezifische Ineffizienz identifiziert werden kann:

$$i_j \sim N(\mu_i, \sigma_i) - \iota. \quad (4.3)$$

Die gruppenspezifische Ineffizienz ι ist ein Verschiebungsparameter, der die Effizienzverteilung i nach links, d.h. in Richtung null verschiebt. In Abbildung 4.2 sind zwei normalverteilte Kurven dargestellt, wobei die Kurve A um die Konstante ι nach links verschoben wurde.

Die Erkenntnisse des dritten Kapitels lassen vermuten, dass das Validitätsmass Δ bei gruppenspezifisch unabhängigen Abschnitten der GPM klein ausfallen sollte. Daher zielen die folgenden Parameter darauf ab, diese Unabhängigkeit zu beeinflussen. Mit der gewählten Transformationsfunktion lassen sich solche unterschiedliche Abschnitte ausschliesslich durch unterschiedliche Inputkombinationen erreichen. Die Erwartungswerte (μ_x^A und μ_x^B) der Inputkombinationen beider Gruppen A und B sind folglich Parameter, die diese Unabhängigkeit der effizienten Abschnitte beeinflussen sollten. Je unterschiedlicher die beiden Erwartungswerte sind, desto kleiner sollte das Validitätsmass Δ ausfallen (in den Abbildungen 3.1 und 3.2 sind die Erwartungswerte der Inputkombinationen beider Gruppen vergleichbar, in Abbildung 3.3 unterscheiden sie sich jedoch). Zusammen mit den Erwartungswerten sind aber auch die Standardabweichungen der Inputkombinationen (σ_x^A und σ_x^B) entscheidend dafür verantwortlich, ob gruppenspezifisch unabhängige Abschnitte definiert werden können: Grosse Standardabweichungen bedeuten, dass die Inputmixe relativ stark um den Erwartungswert streuen und die Bereiche der realisierten Inputmixe *ceteris paribus* tendenziell stärker überlappend sind. Grosse Standardabweichungen sollten daher *ceteris paribus* mit einem hohen Wert des Validitätsmasses Δ einhergehen.⁵ Schliesslich wird noch die Standardabweichung σ_i der stochastischen Effizienzverteilung als Parameter variiert. Dieser Parameter soll kontrollieren, ob die Resultate über einen weiten Bereich möglicher Effizienzverteilungen Gültigkeit haben.

4.3 Phase 2: Berechnung der Daten

Die Zufallszahlen werden mit MATLAB generiert.⁶ Einige Parameter zur Berechnung der Daten wurden während der Simulation nicht verändert. Zu diesen Parametern gehört die Stichprobengrösse, die auf 200 Beobachtungen fixiert ist, wobei beide Gruppen gleich stark vertreten sind. Die Grösse der Stichprobe in einer Simulation sollte sich nach den in

⁵Sind Erwartungswerte und Standardabweichungen für die Outputmengen und die Inputmixe für beide Gruppen identisch, sind die Beobachtungen in der Technologiemenge i.i.d., d.h. unabhängig und identisch verteilt. In diesen Parameterkombinationen ist die implizite Konsistenzbedingung erfüllt und die Effizienzdifferenzen müssten unverzerrt und valide geschätzt werden können.

⁶MATLAB 6.1 verwendet die sogenannte *Ziggurat*-Methode zur Generierung normalverteilter Zufallszahlen [vgl. Marsaglia und Tsang (1984)]. Die Leistungsfähigkeit dieser Methode lässt sich anhand eines Beispiels illustrieren: Mit einem mit 800MHz getakteten Computer lassen sich pro Sekunde 10 Millionen normalverteilte Zufallszahlen generieren. Ein solcher Computer müsste während 58 000 Jahren ununterbrochen Zufallszahlen ziehen, bevor sie sich wiederholen.

Tabelle 4.1: Parameterwerte der Simulation

| Parameter | Gruppe A | Gruppe B |
|---|----------|----------|
| Standardabweichung Effizienzverteilung (σ_i) | 0.15 | 0.15 |
| | 0.20 | 0.20 |
| | 0.25 | 0.25 |
| Gruppenspezifische Ineffizienz (ι) | 0.00 | 0.00 |
| | 0.05 | 0.00 |
| | 0.10 | 0.00 |
| | 0.15 | 0.00 |
| | 0.20 | 0.00 |
| Erwartungswert Inputmix (μ_x) | 50 | 50 |
| | 56 | 44 |
| | 62 | 38 |
| | 68 | 32 |
| | 74 | 26 |
| | 80 | 20 |
| Standardabweichung Inputmix (σ_x) | 10 | 10 |
| | 7 | 7 |
| | 4 | 4 |

der empirischen Forschung relevanten bzw. üblichen Stichprobengrößen richten. Banker (1996) simuliert z.B. vier Stichprobengrößen von 50 bis 200 Beobachtungen. Kittelsen (1999) dagegen simuliert zum überwiegenden Teil Stichprobengrößen von 100 Beobachtungen, wobei in einem Teil der Simulation die Anzahl der Beobachtungen auf 1000 erhöht wird. Der erwartete Output beider Gruppen beträgt je 100 Einheiten mit einer Standardabweichung von 20 Einheiten.⁷ Die erwartete Effizienz beider Gruppen wurde auf $\mu_i = 0.85$ festgelegt. Beide Gruppen sollten demzufolge im Erwartungswert 15% ihrer eingesetzten Inputs einsparen können bzw. verwenden rund 17.65% zu viel Inputs .

Eine Übersicht über die Parameterwerte, die während der Simulation verändert werden, ist in Tabelle 4.1 aufgeführt. Die Parameterwerte sind zeilenweise für die beiden Gruppen aufgeführt. Für die Standardabweichung der zu erwartenden Effizienz werden

⁷Bei diesen Verteilungsannahmen ist ein nichtpositiver Output sehr unwahrscheinlich. Sollte dies trotzdem geschehen, wird der Output gleich eins gesetzt.

Werte von 0.15, 0.20 und 0.25 simuliert, wobei dieser Parameter für beide Gruppen immer denselben Wert annimmt (vgl. die ersten drei Zeilen der Tabelle 4.1). Damit sichergestellt ist, dass die Transformationsfunktion in Gleichung 4.1 ausschliesslich nicht-negative Werte annimmt, werden realisierte Werte für i_j , die grösser als eins sind, gleich eins gesetzt. Die Wahrscheinlichkeiten, dass ein solches Ereignis eintritt beträgt für $\sigma_i = 0.15$ rund 0.159, für $\sigma_i = 0.20$ über 0.227 und für $\sigma_i = 0.25$ über 0.274. Wird dagegen ein nicht-positiver Wert i_j für eine Beobachtung j realisiert (was negative Inputmengen nach sich ziehen würde), wird dieser auf den Erwartungswert minus die gruppenspezifische Ineffizienz (d.h. $\mu_i - \iota$) gesetzt, den modalen Wert der entsprechenden Effizienzverteilung. Ein solches Ereignis tritt für $\sigma_i = 0.15$ *de facto* nie auf. Für $\sigma_i = 0.20$ ist ein solches Ereignis ein Mal in einer Million Fälle und für $\sigma_i = 0.25$ schon ein Mal in 2700 Fällen zu erwarten (diese Wahrscheinlichkeitsangaben basieren auf einem $\iota = 0$).

In Bezug auf die Hypothesentests ist relevant, dass die Effizienzverteilung beider Gruppen gleich ist. Die gruppenspezifische Ineffizienz ι^A , die Werte von 0, 0.05, 0.10, 0.15 und 0.20 annimmt, verschiebt die Effizienzverteilung der Gruppe A lediglich nach links (d.h. in Richtung null, vgl. Abbildung 4.2), während ι^B immer gleich null ist. Somit resultiert z.B. für $\iota^A = 0.20$ eine erwartete Effizienz für die Gruppe A von $i^A = 0.65$ ($= 0.85 - 0.20$) mit einer Obergrenze von 0.80 ($= 1.00 - 0.20$).

Der Inputmix μ_x , der die Unterschiede in der Technologie repräsentiert, nimmt für die Gruppe A, ausgehend von einem Erwartungswert von 50 Prozent, in Schritten von 6 Prozentpunkten bis auf 80 Prozent zu. Gleichzeitig wird der Inputmix μ_x^B der Gruppe B von 50 Prozent mit denselben Inkrementen bis auf 20 Prozent gesenkt. In der Ausgangslage ist der erwartete Anteil des ersten Inputs an der Summe der Inputs bei beiden Gruppen bei 50 Prozent; beide Gruppen verwenden dieselbe Technologie. Dann werden Inputmixe von 56 (Gruppe A) und 44 (Gruppe B) Prozent miteinander verglichen (vgl. zehnte Zeile der Tabelle 4.1). In den Parameterkombinationen mit den unterschiedlichsten Technologien werden Inputkombinationen von 80 und 20 Prozent simuliert. Die Standardabweichung σ_x , die die Streuung der erwarteten Inputmixe definiert, nimmt Werte von 10, 7 und 4 an – wieder sind die Werte jeweils für beide Gruppen identisch.

Jede der 270 ($= 3 * 5 * 6 * 3$) möglichen Parameterkombinationen wird dabei in 500 Iterationen simuliert (insgesamt 135 000 simulierte Effizienzmessungen). In jeder Iteration werden für die 200 generierten Beobachtungen eine DEA zur Schätzung der Effizienz und zusätzlich zwei für die Analyse der Validität durchgeführt (vgl. Abschnitt 4.4); pro Iteration werden folglich 600 Simplexalgorithmen berechnet. Im Ganzen werden somit rund 81 Millionen Simplexalgorithmen gelöst.

4.4 Phase 3: DEA, Analyse der Validität und Hypothesentests

Die berechneten Input-Outputdaten werden anschliessend in einer inputorientierten DEA mit konstanten Skalenerträgen analysiert (vgl. Gleichung 2.15 bzw. 2.19). Konstante Skalenerträge sind notwendig, da die Transformationsfunktion des Datengenerierungsprozess ebenfalls konstante Skalenerträge aufweist (die Summe der beiden Exponenten in Gleichung 4.1 ist gleich eins).⁸ Ein Vorteil der DEA-Version mit konstanten Skalenerträgen besteht darin, dass sich die geschätzte Effizienz für die input- und outputorientierte DEA nicht unterscheidet. Folglich sind auch die Resultate der Simulation nicht von der Orientierung der Ineffizienz abhängig.

Das Ziel der anschliessenden Analyse der Validität ist die Ermittlung des Validitätsmasses Δ . Dazu werden die ursprünglichen Inputdaten der Gruppe A zuerst mit dem Faktor $\Psi = 1.\overline{11}$ multipliziert und die DEA wiederholt.⁹ Das Validitätsmass $\Delta_{\Psi=1.\overline{11}}$ für Veränderung der Daten in Richtung „ineffizienter“ wird gemäss Gleichungen 3.21, 3.22 und 3.23 berechnet. Um das Validitätsmass auch in Richtung „effizienter“ zu bestimmen, wird dieselbe Prozedur mit dem Faktor $\Psi = 0.\overline{90}$ wiederholt und $\Delta_{\Psi=0.\overline{90}}$ berechnet.¹⁰ Damit das Validitätsmass Δ die Validität in beide Richtungen erfasst, wird das arithmetischen Mittel von $\Delta_{\Psi=1.\overline{11}}$ und $\Delta_{\Psi=0.\overline{90}}$ gebildet.¹¹

⁸Die Strukturannahmen der DEA müssen sich nach der zugrunde liegenden Technologie richten. Kommt eine DEA-Version mit konstanten Skalenerträgen bei einer Technologie, die durch variable Skalenerträge gekennzeichnet ist, zur Anwendung, sind die geschätzten Effizienzwerte nicht konsistent.

⁹Dieser Wert von Ψ führt zu einer Reduktion der wahren Effizienz der Gruppe A um 10%

¹⁰Dieser Wert von Ψ führt zu einer Zunahme der wahren Effizienz der Gruppe A um 10%

¹¹Grundsätzlich sollten $\Delta_{\Psi=1.\overline{11}}$ und $\Delta_{\Psi=0.\overline{90}}$ identisch sein. Die beiden Werte können sich aber z.B. dann unterscheiden, wenn die GPM aus gruppenspezifisch unabhängigen Abschnitten besteht, bei der die Beobachtungen der Gruppe A nach der Ausdehnung der Inputs von Beobachtungen der Gruppe B

Zudem wird in dieser Phase geprüft, ob die geschätzten Effizienzdifferenzen zwischen den Gruppen statistisch signifikant von null verschieden sind. Für die statistischen Tests gelten folgende Hypothesen:

- $H_0: \theta^A = \theta^B$
- $H_A: \theta^A \neq \theta^B$.

Ein Fehler erster Art (α -Fehler) liegt vor, wenn

- die Nullhypothese ($\theta^A = \theta^B$) irrtümlich abgelehnt wird, obwohl diese der wahren Situation entspricht. Dieser Fehler tritt auf, wenn die Hypothesentests $\theta^A \neq \theta^B$ ergeben, obwohl die Daten mit den Parametern $\iota^A = \iota^B = 0$ generiert wurden.

Ein Fehler zweiter Art (β -Fehler) liegt dagegen vor, wenn

- die Nullhypothese ($\theta^A = \theta^B$) irrtümlich nicht verworfen wird, obwohl diese nicht der wahren Situation entspricht. Dieser Fehler tritt auf, wenn gemäss Hypothesentest die H_0 nicht abgewiesen werden kann, obwohl die Daten mit den Parametern $\iota^A > \iota^B$ generiert wurden.

Die wahre Situation wird durch den Parameter der gruppenspezifischen Ineffizienz ι definiert. Ist ι gleich null, sind beide Gruppen gleich effizient. Ist dagegen ι grösser null, ist die Gruppe A weniger effizient. Die Wahrscheinlichkeit der Fehler erster Art hängt ausschliesslich vom Signifikanzniveau ab. Die Wahrscheinlichkeit eines Fehlers zweiter Art ist zusätzlich abhängig von den Parametern μ_i , σ_i und ι , d.h. von der Verteilung der Effizienz und vom Ausmass der Linksverschiebung der Effizienzverteilung der Gruppe A. Da die Effizienzverteilungen beider Gruppen normalverteilt sind, ist auch die Differenz dieser Effizienzverteilungen normalverteilt mit dem Erwartungswert $\mu_i^D = \mu_i^B - (\mu_i^A - \iota)$ und der Standardabweichung $\sigma_i^D = \sigma_i \sqrt{\frac{1}{N_A} + \frac{1}{N_B}}$. Da in allen simulierten Szenarien $\mu_i^B = \mu_i^A$ gilt, ist μ_i^D gleich ι . In Tabelle 4.2 sind die erwarteten Anteile der Fehler erster bzw. zweiter Art für die simulierten Parameterkombinationen und ein Signifikanzniveau von 95% auf-

dominiert werden. Dann ist $\Delta_{\Psi=1.11}$ grösser als $\Delta_{\Psi=0.90}$, da die Beobachtungen nach der Ausdehnung der Inputs dominiert und somit ineffizienter werden, während bei einer Reduktion der Inputs diese aber nicht effizienter werden.

Tabelle 4.2: Erwartete Fehlerwahrscheinlichkeiten der Hypothesentests

| $\mu_i^D = \iota^A$ | $\sigma_i^D = 0.15$ | $\sigma_i^D = 0.20$ | $\sigma_i^D = 0.25$ |
|---------------------|---------------------|---------------------|---------------------|
| 0.00 | 0.0500 | 0.0500 | 0.0500 |
| 0.05 | 0.3457 | 0.5762 | 0.7074 |
| 0.10 | 0.0029 | 0.0576 | 0.1926 |
| 0.15 | 0.0000 | 0.0004 | 0.0112 |
| 0.20 | 0.0000 | 0.0000 | 0.0001 |

geführt.¹² Da ausschliesslich in Situationen, in denen beide Gruppen gleich effizient sind, Fehler erster Art begangen werden können, sind diese Wahrscheinlichkeiten in der ersten Zeile der Tabelle 4.2 aufgeführt ($\mu_i^D = \iota^A = 0$). Sind die Effizienzdifferenzen valide, ist bei den Hypothesentests mit gleich effizienten Gruppen mit einem Anteil von fünf Prozent falscher Aussagen zu rechnen. In den Zeilen zwei bis fünf sind die Wahrscheinlichkeiten der Fehler zweiter Art aufgeführt. Je grösser die Streuung und je kleiner der Erwartungswert der Effizienzdifferenz ist, desto grösser ist die Fehlerwahrscheinlichkeit. Aus der letzten Spalte der zweiten Zeile ist zu entnehmen, dass in über 70% aller Hypothesentests mit den entsprechenden Parameterwerten σ_i^D und ι die Nullhypothese zu unrecht nicht abgewiesen werden müsste. Mit zunehmenden erwarteten Effizienzdifferenzen sinkt dieser Anteil unter ein Promille.¹³

¹²Die in der Tabelle 4.2 ausgewiesenen erwarteten Anteile der Fehler zweiter Art (Zeilen zwei bis fünf) basieren auf der Annahme, dass die geschätzten Effizienzdifferenzen unverzerrt und somit valide sind. Die Tabelle 4.2 dient aber auch bei verzerrt geschätzten Effizienzdifferenzen als Referenz. Der Anteil der Fehler zweiter Art hängt von der effektiv realisierten Effizienzdifferenz ab, die im Gegensatz zum Verschiebungsparameter ι um die Verzerrung reduziert ist. Wird die erste Spalte als realisierte Effizienzdifferenz interpretiert, lässt sich der erwartete Anteil der Fehler zweiter Art in Abhängigkeit von σ_i direkt ablesen. Angenommen, dass die wahre Effizienzdifferenz 20% beträgt, aber eine geschätzte erwartete Effizienzdifferenz von 5% realisiert wird, ist der Tabelle zu entnehmen, dass der Anteil der Fehler zweiter Art für $\sigma_i = 0.25$ über 70% betragen sollte (vgl. zweite Zeile, letzte Spalte).

¹³In Tabelle B.1 im Anhang B sind die erwarteten Anteile der Fehler erster und zweiter Art für realisierte Effizienzdifferenzen zwischen null und 20 Prozent aufgeführt.

4.5 Phase 4: Auswertung der Simulation

4.5.1 Geschätzte erwartete Verzerrungen

In diesem Abschnitt soll die Frage geklärt werden, wie stark die geschätzte Effizienz tatsächlich verzerrt ist. Die geschätzte erwartete Verzerrung \mathcal{B} der geschätzten Effizienz einer Gruppe von Beobachtungen entspricht der Differenz zwischen dem geschätzten Erwartungswert der Effizienz ($\bar{\bar{\theta}}$) und dem Erwartungswert der wahren Effizienz $E(\theta)$ [vgl. Kennedy (1998), S. 24]:

$$\begin{aligned}\mathcal{B} &= \bar{\bar{\theta}} - E(\theta) \\ \bar{\bar{\theta}} &= \sum_{i=1}^I \bar{\theta}_i / I \\ \bar{\theta}_i &= \sum_{j=1}^N \hat{\theta}_{ij} / N,\end{aligned}$$

wobei $\bar{\theta}_i$ die geschätzte durchschnittliche DEA-Effizienz in der Iteration i und N die Anzahl Beobachtungen in jeder der I Iterationen ist.

Wie aus Tabelle 4.3 ersichtlich ist, sind die geschätzten erwarteten Verzerrungen der Gruppe A primär auf den Parameter Inputmix μ_x zurückzuführen. Je unterschiedlicher die Technologie in Bezug auf den Inputmix der beiden Gruppen ist, desto grösser sind – bei gegebener Effizienzdifferenz ι^A – die geschätzten erwarteten Verzerrungen (vgl. Tabelle 4.3). Die Standardabweichung der Inputmixe σ_x hat kaum einen Einfluss auf die minimal bzw. maximal realisierten Werte der geschätzten erwarteten Verzerrung. Die minimalen Verzerrungen für $\iota = 0.20$ liegen für identische Inputmixe zwischen 0.066 und 0.021 Prozentpunkten, während die maximalen Verzerrungen Werte zwischen 15.135 und 15.554 Prozentpunkten annehmen (vgl. letzte Spalte). Die Unterschiede, die auf die Standardabweichung zurückzuführen sind, können bei den extremen Inputmixin somit als vernachlässigbar bezeichnet werden. Allerdings nehmen die Verzerrungen bei kleinerer Standardabweichung mit zunehmenden Unterschieden bei den Inputmixin schneller zu als bei den grösseren Standardabweichungen. In Tabelle 4.3 ist dies z.B. in der vierten bis sechsten Zeile gut ersichtlich. Bei einer Standardabweichung von 10% nimmt die geschätzte erwartete Verzerrung für $\iota^A = 0.20$ einen Wert von 1.3% an, während bei einer Standardabweichung von 4% bereits über 5 Prozent Verzerrung realisiert wird.

In einer empirischen Anwendung sind diese Parameterwerte i.d.R. nicht bekannt. Da untersucht werden soll, ob das Ausmass der Verzerrung durch das Validitätsmass Δ erklärt werden kann, sind in Tabelle 4.4 die geschätzten erwarteten Verzerrungen der Grup-

Tabelle 4.3: Geschätzte erwartete Verzerrungen der Gruppe A (\mathcal{B}^A), Teil 1^{*)}

| Inputmix μ_x | Std.abw. σ_x | $\iota^A = 0$ | $\iota^A = 0.05$ | $\iota^A = 0.10$ | $\iota^A = 0.15$ | $\iota^A = 0.20$ |
|------------------|---------------------|---------------|------------------|------------------|------------------|------------------|
| 50:50 | 10 | 0.003 | 0.076 | 0.092 | 0.111 | 0.066 |
| | 7 | 0.005 | 0.054 | 0.071 | 0.067 | 0.041 |
| | 4 | 0.002 | 0.044 | 0.042 | 0.038 | 0.021 |
| 56:44 | 10 | 0.009 | 0.578 | 0.916 | 1.197 | 1.301 |
| | 7 | 0.004 | 0.947 | 1.497 | 1.937 | 2.254 |
| | 4 | 0.002 | 1.903 | 3.265 | 4.382 | 5.235 |
| 62:38 | 10 | 0.012 | 1.929 | 0.003 | 4.152 | 4.899 |
| | 7 | 0.003 | 2.878 | 4.893 | 6.358 | 7.518 |
| | 4 | 0.003 | 3.896 | 6.939 | 9.289 | 11.296 |
| 68:32 | 10 | 0.013 | 3.320 | 5.986 | 8.034 | 9.694 |
| | 7 | 0.004 | 3.926 | 7.432 | 10.291 | 12.623 |
| | 4 | 0.011 | 4.041 | 7.975 | 11.649 | 14.796 |
| 74:26 | 10 | 0.014 | 3.887 | 7.537 | 10.835 | 13.563 |
| | 7 | 0.003 | 4.026 | 7.953 | 11.669 | 15.099 |
| | 4 | 0.005 | 4.056 | 7.998 | 11.844 | 15.527 |
| 80:20 | 10 | 0.002 | 4.006 | 7.941 | 11.696 | 15.135 |
| | 7 | 0.014 | 4.041 | 8.008 | 11.845 | 15.559 |
| | 4 | 0.009 | 4.061 | 8.036 | 11.867 | 15.554 |

*) Zur übersichtlicheren Darstellung wird die Effizienz θ in Prozentwerte umgerechnet und ist somit auf einem Bereich zwischen null und 100 definiert. Die ausgewiesenen Verzerrungen sind als Prozentpunkte definiert.

pe A zeilenweise für zehn Wertebereiche von Δ ausgewiesen.¹⁴ Für die Beobachtungen der Gruppe A ergibt sich folgendes Bild. Bei den Parameterkombinationen ohne Effizienzunterschiede ($\iota^A = 0$, vgl. erste Spalte) sind die geschätzten erwarteten Verzerrungen kleiner als ein Prozentpunkt¹⁵, wobei das Ausmass negativ mit dem Wert von Δ zusammenhängt (vgl. zweite Spalte). Aus der zehnten Zeile wird zudem ersichtlich, dass die Verzerrun-

¹⁴Da das Validitätsmass Δ auf einem stetigen Wertebereich definiert ist, wird es, zur übersichtlicheren Darstellung, in zehn Kategorien eingeteilt.

¹⁵Zur übersichtlicheren Darstellung wird die Effizienz θ in Prozentwerte umgerechnet und ist somit auf einem Bereich zwischen null und 100 definiert. Die ausgewiesenen Verzerrungen sind als Prozentpunkte definiert.

Tabelle 4.4: Geschätzte erwartete Verzerrungen der Gruppe A (\mathcal{B}^A), Teil 2^{*)}

| Delta Δ | $\iota^A = 0.00$ | $\iota^A = 0.05$ | $\iota^A = 0.10$ | $\iota^A = 0.15$ | $\iota^A = 0.20$ |
|----------------|------------------|------------------|------------------|------------------|------------------|
| 0 bis 0.1 | 0.279 | 4.335 | 8.283 | 12.121 | 15.815 |
| über 0.1 - 0.2 | 0.196 | 4.091 | 7.833 | 11.568 | 15.138 |
| über 0.2 - 0.3 | 0.249 | 3.786 | 7.179 | 10.647 | 14.069 |
| über 0.3 - 0.4 | 0.214 | 3.360 | 6.445 | 9.572 | 12.514 |
| über 0.4 - 0.5 | 0.200 | 2.863 | 5.550 | 8.417 | 11.115 |
| über 0.5 - 0.6 | 0.159 | 2.308 | 4.454 | 6.966 | 9.320 |
| über 0.6 - 0.7 | 0.153 | 1.840 | 3.555 | 5.480 | 7.342 |
| über 0.7 - 0.8 | 0.170 | 1.366 | 2.627 | 4.126 | 5.656 |
| über 0.8 - 0.9 | 0.170 | 0.878 | 1.686 | 2.599 | 3.758 |
| über 0.9 | 0.133 | 0.310 | 0.435 | 0.594 | 0.816 |

^{*)} Zur übersichtlicheren Darstellung wird die Effizienz θ in Prozentwerte umgerechnet und ist somit auf einem Bereich zwischen null und 100 definiert. Die ausgewiesenen Verzerrungen sind als Prozentpunkte definiert.

gen für Werte von $\Delta > 0.9$ unabhängig vom Verschiebungsparameter ι^A kleiner als ein Prozentpunkt ausfallen. Allgemein kann festgehalten werden, dass die Verzerrungen bei gegebener gruppenspezifischer Ineffizienz ι^A mit abnehmenden Werten von Δ zunehmen. Interessant ist der Anteil der gruppenspezifischen Ineffizienz ι^A , der durch die Verzerrungen überlagert wird. In der fünften Spalte, bei der die gruppenspezifische Ineffizienz 20% beträgt, wird für $\Delta < 0.1$ rund 79% der wahren Effizienzunterschiede durch Verzerrungen überlagert ($79 = 15.815/20$). Der Anteil der Effizienzunterschiede, der durch Verzerrungen verdeckt wird, steht unabhängig vom Ausmass der Effizienzunterschiede in einem negativen Zusammenhang mit dem Wert des Validitätsmasses Δ .

Folgerung 4.1 *Existieren keine Effizienzunterschiede zwischen den beiden Gruppen, sind die Verzerrungen der geschätzten Effizienz der Gruppe A unabhängig vom Validitätsmass kleiner als ein Prozentpunkt.¹⁶ Existieren Effizienzunterschiede, sind die Verzerrungen nur für Werte des Validitätsmasses Δ im Bereich um eins vernachlässigbar klein.¹⁷*

¹⁶Da die Wahrscheinlichkeit einer vollständig effizienten Beobachtung in der Simulation für Parameterkombinationen mit $\iota^A = 0$ bei beiden Gruppen gleich gross ist, ist die erweiterte explizite Konsistenzbedingung erfüllt und die Effizienz sollte unabhängig der impliziten Konsistenzbedingung nahezu unverzerrt sein (vgl. Abschnitt 3.3.4).

¹⁷In der Simulation ist bei Parameterkombinationen mit $\iota^A > 0$ die erweiterte Konsistenzbedingung

Tabelle 4.5: Geschätzte erwartete Verzerrungen der Gruppe B (\mathcal{B}^B)^{*)}

| Delta Δ | $\iota^A = 0.00$ | $\iota^A = 0.05$ | $\iota^A = 0.10$ | $\iota^A = 0.15$ | $\iota^A = 0.20$ |
|----------------|------------------|------------------|------------------|------------------|------------------|
| 0 bis 0.1 | 0.277 | 0.312 | 0.317 | 0.330 | 0.346 |
| über 0.1 - 0.2 | 0.195 | 0.236 | 0.313 | 0.336 | 0.328 |
| über 0.2 - 0.3 | 0.266 | 0.281 | 0.246 | 0.313 | 0.322 |
| über 0.3 - 0.4 | 0.213 | 0.260 | 0.280 | 0.245 | 0.276 |
| über 0.4 - 0.5 | 0.192 | 0.234 | 0.288 | 0.298 | 0.272 |
| über 0.5 - 0.6 | 0.165 | 0.211 | 0.247 | 0.284 | 0.309 |
| über 0.6 - 0.7 | 0.142 | 0.205 | 0.213 | 0.251 | 0.303 |
| über 0.7 - 0.8 | 0.171 | 0.234 | 0.254 | 0.246 | 0.274 |
| über 0.8 - 0.9 | 0.167 | 0.234 | 0.260 | 0.274 | 0.267 |
| über 0.9 | 0.135 | 0.226 | 0.253 | 0.261 | 0.267 |

^{*)} Zur übersichtlicheren Darstellung wird die Effizienz θ in Prozentwerte umgerechnet und ist somit auf einem Bereich zwischen null und 100 definiert. Die ausgewiesenen Verzerrungen sind als Prozentpunkte definiert.

In Tabelle 4.5 sind die geschätzten erwarteten Verzerrungen der Beobachtungen der Gruppe B aufgeführt. Die Verzerrungen sind dabei durchwegs kleiner als ein halber Prozentpunkt. Diese geringen Verzerrungen sind ein Indiz für die Güte der DEA-Schätzer unter vorteilhaften Umständen.¹⁸ Trotzdem ist zwischen den Verzerrungen und dem Δ ein negativer Zusammenhang festzustellen (vgl. z.B. die erste und letzte Zeile der ersten Spalte).

Folgerung 4.2 *Unabhängig von den wahren Effizienzdifferenzen (ι^A) und vom Validitätsmass sind die Verzerrungen der geschätzten Effizienz der Gruppe B kleiner als ein halber Prozentpunkt. Es ist allerdings ein Zusammenhang zwischen den Verzerrungen und dem Validitätsmass bzw. den wahren Effizienzunterschieden (ι^A) festzustellen, der aber nur für eine Differenz von 0.25 Prozentpunkten verantwortlich ist.*

Wie in Abschnitt 3.1 argumentiert wurde, sind für die Hypothesentests nicht unverzerrt verletzt, da die Wahrscheinlichkeit einer vollständig effizienten Beobachtung bei der Gruppe A gleich null ist. Somit muss für unverzerrte Effizienzdifferenzen die implizite Konsistenzbedingung zwingend erfüllt sein, was – wie die Simulation nahe legt – nur bei hohen Werten des Validitätsmasses gegeben ist.

¹⁸Die Wahrscheinlichkeit einer vollständig effizienten Beobachtung ist bei der Gruppe B in allen Parameterkombinationen immer grösser null. Zudem sind die Beobachtungen der Gruppe B – isoliert betrachtet – i.i. in der Technologiemenge verteilt.

Tabelle 4.6: Die geschätzten erwarteten Verzerrungen der Effizienzdifferenzen \mathcal{B}^{AB*})

| Delta Δ | $\iota^A = 0.00$ | $\iota^A = 0.05$ | $\iota^A = 0.10$ | $\iota^A = 0.15$ | $\iota^A = 0.20$ |
|----------------|------------------|------------------|------------------|------------------|------------------|
| 0 bis 0.1 | 0.003 | 4.024 | 7.966 | 11.791 | 15.469 |
| über 0.1 - 0.2 | 0.000 | 3.856 | 7.520 | 11.232 | 14.810 |
| über 0.2 - 0.3 | -0.017 | 3.506 | 6.934 | 10.334 | 13.747 |
| über 0.3 - 0.4 | 0.001 | 3.100 | 6.165 | 9.328 | 12.238 |
| über 0.4 - 0.5 | 0.008 | 2.629 | 5.263 | 8.119 | 10.843 |
| über 0.5 - 0.6 | -0.006 | 2.097 | 4.207 | 6.681 | 9.010 |
| über 0.6 - 0.7 | 0.011 | 1.635 | 3.343 | 5.228 | 7.039 |
| über 0.7 - 0.8 | -0.001 | 1.132 | 2.373 | 3.880 | 5.382 |
| über 0.8 - 0.9 | 0.003 | 0.645 | 1.426 | 2.324 | 3.490 |
| über 0.9 | -0.001 | 0.084 | 0.182 | 0.333 | 0.549 |

*) Zur übersichtlicheren Darstellung wird die Effizienz θ in Prozentwerte umgerechnet und ist somit auf einem Bereich zwischen null und 100 definiert. Die ausgewiesenen Verzerrungen sind als Prozentpunkte definiert.

te Schätzer der Effizienz, sondern unverzerrte, d.h. valide Effizienzdifferenzen relevant. Die geschätzte erwartete Verzerrung der Effizienzdifferenz ist gleich $\mathcal{B}^{AB} = \mathcal{B}^A - \mathcal{B}^B$. In Tabelle 4.6 sind die geschätzten erwarteten Verzerrungen der Effizienzdifferenzen aufgeführt. Existieren keine Effizienzunterschiede (die erweiterte explizite Konsistenzbedingung ist bei diesen Parameterkombinationen erfüllt), sind die geschätzten Effizienzdifferenzen sehr valide: Die geschätzten erwarteten Verzerrungen betragen im Absolutwert weniger als 0.017 Prozentpunkte. Beim Vorliegen von Effizienzdifferenzen ist die erweiterte explizite Konsistenzbedingung verletzt. Für Werte von Δ über 0.9 sind die geschätzten erwarteten Verzerrungen der Effizienzdifferenz jedoch durchwegs kleiner als ein Prozentpunkt (vgl. letzte Zeile). Die implizite Konsistenzbedingung scheint in diesen Parameterkombinationen erfüllt zu sein und stellt sicher, dass die Effizienzdifferenzen nahezu valide sind. In allen anderen Fällen sind weder die erweiterte explizite noch die implizite Konsistenzbedingung erfüllt, die Effizienzdifferenzen sind folglich weniger valide und es ist mit einer Zunahme der Wahrscheinlichkeit von Fehlern zweiter Art bei den Hypothesentests zu rechnen.

Folgerung 4.3 *Existieren keine Effizienzunterschiede zwischen den beiden Gruppen, ist in der Simulation die erweiterte explizite Konsistenzbedingung erfüllt: Die geschätzten er-*

warteten Verzerrungen der Effizienzdifferenzen sind sehr klein und dürfen – unabhängig von der impliziten Konsistenzbedingung – als valide bezeichnet werden. Ist die erweiterte explizite Konsistenzbedingung verletzt ($\iota^A > 0$), sind für Werte von Δ nahe eins die geschätzten erwarteten Verzerrungen der Effizienzdifferenzen durchwegs kleiner als ein Prozentpunkt. Dies ist ein Indiz dafür, dass die implizite Konsistenzbedingung erfüllt ist. In allen anderen Parameterkombinationen (weder die erweiterte explizite noch die implizite Konsistenzbedingung ist erfüllt) sind die geschätzten erwarteten Verzerrungen der Effizienzdifferenzen beachtlich und stehen in einem negativen Zusammenhang mit Validitätsmasses Δ .

In Tabelle 4.7 ist die theoretische Validität der geschätzten Effizienz aufgeführt. Die theoretische Validität ist definiert als der Korrelationskoeffizient zwischen der wahren und der geschätzten Effizienz [vgl. Schnell, Hill und Esser (1999), Kap. 4]. Werte von eins deuten auf eine hohe theoretische Validität hin. In der zweiten und dritten Spalte ist die theoretische Validität für die wahre Situation „beide sind gleich effizient“ aufgeführt, welche für beide Gruppen unabhängig vom Validitätsmass Δ nur knapp unter dem idealen Wert von eins liegt. Die vierte und fünfte Spalte enthält die theoretische Validität für die Situationen, in denen die Gruppe A weniger effizient ist. Deutlich wird, dass die theoretische Validität für kleine Werte von Δ geringer ausfällt, aber kontinuierlich mit Δ zunimmt. Die theoretische Validität der geschätzten Effizienz der Gruppe B ist, gemäss Spalte fünf, unabhängig vom Validitätsmass nahe bei eins.

Der Zusammenhang zwischen dem Validitätsmass und der geschätzten erwarteten Verzerrung der Effizienzdifferenz wird anhand der folgenden Gleichung in einer linearen Kurvenanpassung ermittelt:

$$\mathcal{B}^{AB}/\iota = a_0 + a_1\Delta. \quad (4.4)$$

Da die geschätzte erwartete Verzerrung der Effizienzdifferenz mit der wahren Effizienzdifferenz (ι) normiert wird, nimmt die abhängige Variable Werte zwischen null und eins an. Für die Konstante a_0 wird ein Wert von 0.84, für a_1 ein Wert von -0.81 ermittelt. Über 97 Prozent der Varianz wird durch diese Gleichung erklärt.

Zusammenfassend kann gesagt werden, dass hohe Werte von Δ unabhängig von den „wahren“ Effizienzunterschieden mit validen Effizienzdifferenzen und sehr hoher theoretischer Validität in Verbindung zu bringen sind. Werte von Δ um eins sind somit eine hinreichende Bedingung für unverzerrt geschätzte und folglich valide Effizienzdifferenzen.

Tabelle 4.7: Theoretische Validität der DEA-Effizienz

| Delta Δ | $\iota^A = 0$ | | $\iota^A > 0$ | |
|----------------|---------------|----------|---------------|----------|
| | Gruppe A | Gruppe B | Gruppe A | Gruppe B |
| 0 bis 0.1 | 0.992 | 0.992 | 0.745 | 0.991 |
| über 0.1 - 0.2 | 0.994 | 0.995 | 0.786 | 0.991 |
| über 0.2 - 0.3 | 0.994 | 0.994 | 0.822 | 0.992 |
| über 0.3 - 0.4 | 0.995 | 0.994 | 0.873 | 0.993 |
| über 0.4 - 0.5 | 0.995 | 0.996 | 0.915 | 0.993 |
| über 0.5 - 0.6 | 0.996 | 0.996 | 0.941 | 0.994 |
| über 0.6 - 0.7 | 0.996 | 0.997 | 0.962 | 0.994 |
| über 0.7 - 0.8 | 0.995 | 0.996 | 0.974 | 0.994 |
| über 0.8 - 0.9 | 0.996 | 0.996 | 0.984 | 0.994 |
| über 0.9 | 0.998 | 0.997 | 0.994 | 0.994 |

Bei kleinen Werten von Δ besteht dagegen die Gefahr, dass potentielle Effizienzunterschiede zu einem grossen Teil durch Verzerrungen überlagert und somit unterschätzt werden. Die Resultate der Simulation stützen die Hypothese 1 (vgl. S. 70).

Folgerung 4.4 *Die anhand der DEA geschätzten Effizienzwerte sind unter Umständen stark nach oben und die Effizienzdifferenzen folglich nach unten verzerrt. Vor allem für kleine Werte von Δ besteht die Gefahr von verzerrten und folglich nicht validen Effizienzdifferenzen. Nimmt dagegen Δ Werte um eins an, sind die geschätzten erwarteten Effizienzdifferenzen über einen breiten Parameterraum nahezu unverzerrt.*

4.5.2 Hypothesentests

Die Frage stellt sich, welche Konsequenzen die Verzerrungen auf die Hypothesentests haben. Da die Verzerrungen ausschliesslich nach oben ausfallen und somit die Effizienzdifferenzen zwischen den beiden Gruppen vermindern, ist zu erwarten, dass beim Vorliegen von Verzerrungen die Hypothesentests zu systematisch falschen Schlussfolgerungen führen.

In der Tabelle 4.8 sind die Resultate der Hypothesentests (zweiseitiger T-Test für unbekannte, aber identische Varianzen) für die Parameterkombinationen mit $\iota^A = \iota^B = 0$ aufgeführt. Das Validitätsmass Δ wurde zur übersichtlicheren Darstellung wieder in zehn Kategorien eingeteilt. In der ersten Spalte ist die Anzahl der Iterationen I angegeben, die

Tabelle 4.8: Anteil α -Fehler ($\iota^A = 0$)

| Delta Δ | I | Anteil α -Fehler | $\iota^A - \mathcal{B}^{AB}$ |
|----------------|--------|-------------------------|------------------------------|
| 0 bis 0.1 | 12 116 | 5.05 | -0.003 |
| über 0.1 - 0.2 | 1 875 | 3.95 | -0.000 |
| über 0.2 - 0.3 | 1 037 | 4.82 | 0.017 |
| über 0.3 - 0.4 | 906 | 5.74 | -0.001 |
| über 0.4 - 0.5 | 921 | 5.32 | -0.008 |
| über 0.5 - 0.6 | 1 414 | 4.67 | 0.006 |
| über 0.6 - 0.7 | 1 044 | 5.27 | -0.011 |
| über 0.7 - 0.8 | 935 | 5.99 | 0.001 |
| über 0.8 - 0.9 | 1 600 | 6.19 | -0.003 |
| über 0.9 | 5 152 | 5.24 | 0.001 |
| Total | 27 000 | 5.12 | -0.001 |

zu einem Δ in der entsprechenden Kategorie geführt hat. Da in jeder Iteration ein Hypothesentest durchgeführt werden kann, entspricht diese Zahl der Anzahl durchgeführter Hypothesentests im entsprechenden Wertebereich von Δ . In der nächsten Spalte ist der relative Anteil der Fehler erster Art aufgeführt. Unabhängig von Δ beträgt der Anteil der Fehler erster Art rund 5% – der Wert, der bei einem Signifikanzniveau von 95% zu erwarten ist. Die letzte Spalte enthält schliesslich die Erwartungswerte der effektiv realisierten Effizienzdifferenzen (in Prozentpunkten), welche den Differenzen zwischen den wahren Effizienzunterschieden (ι^A) und den geschätzten Verzerrungen der Effizienzdifferenzen (\mathcal{B}^{AB}) entsprechen.¹⁹

Folgerung 4.5 *Sind die Effizienzverteilungen beider Gruppen identisch, liefern die Hypothesentests hervorragende Resultate in dem Sinne, dass der Anteil der Fehler erster Art genau der Fehlerwahrscheinlichkeit der Hypothesentests entspricht.*

Die Ergebnisse der Hypothesentests für $\iota^A > \iota^B = 0$ sind für die vier Parameterwerte von ι^A in den Tabellen 4.9 bis 4.12 aufgeführt. In der ersten Spalte findet sich jeweils die Anzahl der durchgeführten Hypothesentests gefolgt vom relativen Anteil der Fehler zweiter Art. In der dritten Spalte sind die Erwartungswerte der effektiv realisierten Effizienzdifferenzen ($\iota^A - \mathcal{B}^{AB}$) aufgeführt. Diese Werte sind die für die Hypothesentests

¹⁹Die Erwartungswerte der realisierten Effizienzdifferenzen entsprechen dem Teil der wahren Effizienzdifferenzen, die nicht durch Verzerrungen überlagert werden.

relevanten Effizienzdifferenzen. Interessant sind sie deshalb, weil sie für die Prognose des theoretisch zu erwartenden Anteils der Fehler zweiter Art massgebend sind. Bei einer realisierten Effizienzdifferenz $\iota^A - \mathcal{B}^{AB}$ von z.B. 2% ist gemäss Tabelle B.1 (im Anhang B) – in Abhängigkeit des Parameters σ_i – mit einem Fehleranteil von 84 bis 92 Prozent zu rechnen.²⁰

In Tabelle 4.9 sind die Resultate der Hypothesentests für eine wahre Effizienzdifferenz von 5% aufgeführt. Bei dieser Differenz wird bei unverzerrt geschätzten Effizienzdifferenzen ein Anteil Fehler zweiter Art zwischen 35 und 70 Prozent erwartet (vgl. Tabellenfussnote oder Tabelle B.1 im Anhang B). Für realisierte Effizienzdifferenzen von knapp 5 Prozent werden in 40 Prozent der Hypothesentests Fehler begangen (vgl. zehnte Zeile), ein Wert, der im Bereich der theoretisch erwarteten Häufigkeit liegt.²¹ Mit abnehmendem Validitätsmass nimmt der realisierte Fehleranteil kontinuierlich zu und für Werte von $\Delta < 0.5$ liegt der Anteil der Fehler bereits über dem theoretisch erwarteten Wert von 70%. Für $\Delta < 0.1$ wird gar ein Anteil Fehler zweiter Art von knapp 93 Prozent realisiert (vgl. erste Zeile).²²

In Tabelle 4.10 sind die Resultate der Hypothesentests für $\iota^A = 0.10$ aufgeführt. Bei dieser Differenz wird bei unverzerrt geschätzten Effizienzdifferenzen mit einem Anteil Fehler zweiter Art zwischen 0.3 und 20 Prozent gerechnet (vgl. Tabellenfussnote oder Tabelle B.1 im Anhang B). Für $\Delta > 0.9$ wird in gut 2 Prozent aller Hypothesentests ein Fehler zweiter Art begangen (vgl. zehnte Zeile). Wieder steigt der Anteil Fehler zweiter Art mit abnehmendem Δ . Bereits für $\Delta < 0.6$ beträgt der Fehleranteil über 20 Prozent und liegt somit über dem erwarteten Wert für die entsprechenden Parameterkombinationen. Der Anteil der realisierten Fehler zweiter Art steigt für $\Delta < 0.1$ sogar auf über 86 Prozent.

Aus Tabelle 4.11 ist ersichtlich, dass für $\iota^A = 0.15$ bei grossen Werten von Δ der

²⁰Auf eine weitere Unterteilung der Resultate für die drei Parameterwerte der Varianz wurde verzichtet. Die in den Tabellen 4.9 bis 4.12 ausgewiesenen Anteile der Fehler zweiter Art sind somit approximativ Mittelwerte der drei simulierten Varianzen.

²¹Auf die Prognose des Anteils der Fehler zweiter Art wird im Folgenden verzichtet. Zusammenfassend kann gesagt werden, dass die prognostizierten Fehleranteile – wie bei den Fehlern erster Art (vgl. Folgerung 4.5) – gut mit den in der Simulation realisierten Anteilen übereinstimmen.

²²Auf die Diskussion der dritten Spalte wird verzichtet. In Abschnitt 4.5.1 wurde gezeigt, dass die Verzerrungen \mathcal{B}^{AB} in einem negativen Zusammenhang mit dem Validitätsmass stehen. Da die Werte in der dritten Spalte als Differenz zwischen ι und der erwarteten Verzerrung \mathcal{B}^{AB} definiert wurden, steht $\iota^A - \mathcal{B}^{AB}$ in einem positiven Zusammenhang mit Δ . Die dritte Spalte wurde ausschliesslich als Referenz zur Prognose der erwarteten Anteile der Fehler zweiter Art in die Tabellen 4.9 bis 4.12 aufgenommen.

Tabelle 4.9: Anteil β -Fehler ($\iota^A = 0.05$)

| Delta Δ | I | Anteil β -Fehler ^{a)} | $\iota^A - \mathcal{B}^{AB}$ |
|----------------|--------|--------------------------------------|------------------------------|
| 0 bis 0.1 | 11 226 | 92.91 | 0.98 |
| über 0.1 - 0.2 | 2 399 | 92.04 | 1.14 |
| über 0.2 - 0.3 | 1 133 | 89.50 | 1.49 |
| über 0.3 - 0.4 | 938 | 87.42 | 1.90 |
| über 0.4 - 0.5 | 1 071 | 81.61 | 2.37 |
| über 0.5 - 0.6 | 1 336 | 76.65 | 2.90 |
| über 0.6 - 0.7 | 1 174 | 67.63 | 3.37 |
| über 0.7 - 0.8 | 1 068 | 59.55 | 3.87 |
| über 0.8 - 0.9 | 1 606 | 50.69 | 4.36 |
| über 0.9 | 5 049 | 40.60 | 4.92 |
| Total | 27 000 | 76.53 | 2.35 |

^{a)} Der theoretisch erwartete Anteil der Fehler zweiter Art beträgt für $\iota^A = 0.05$ und ein Signifikanzniveau von 95% 34.6 bis 70.7 Prozent.

Tabelle 4.10: Anteil β -Fehler ($\iota^A = 0.10$)

| Delta Δ | I | Anteil β -Fehler ^{a)} | $\iota^A - \mathcal{B}^{AB}$ |
|----------------|--------|--------------------------------------|------------------------------|
| 0 bis 0.1 | 9 953 | 86.10 | 2.03 |
| über 0.1 - 0.2 | 1 860 | 79.09 | 2.48 |
| über 0.2 - 0.3 | 1 952 | 74.08 | 3.07 |
| über 0.3 - 0.4 | 1 079 | 60.15 | 3.84 |
| über 0.4 - 0.5 | 992 | 48.29 | 4.74 |
| über 0.5 - 0.6 | 1 269 | 33.57 | 5.79 |
| über 0.6 - 0.7 | 1 351 | 22.50 | 6.66 |
| über 0.7 - 0.8 | 1 182 | 10.41 | 7.63 |
| über 0.8 - 0.9 | 1 595 | 6.33 | 8.57 |
| über 0.9 | 5 767 | 2.25 | 9.82 |
| Total | 27 000 | 76.53 | 2.35 |

^{a)} Der theoretisch erwartete Anteil der Fehler zweiter Art beträgt für $\iota^A = 0.10$ und ein Signifikanzniveau von 95% 3 bis 19.3 Prozent.

Tabelle 4.11: Anteil β -Fehler ($\iota^A = 0.15$)

| Delta Δ | I | Anteil β -Fehler ^{a)} | $\iota^A - \mathcal{B}^{AB}$ |
|----------------|--------|--------------------------------------|------------------------------|
| 0 bis 0.1 | 8 417 | 75.69 | 3.21 |
| über 0.1 - 0.2 | 1 985 | 64.69 | 3.77 |
| über 0.2 - 0.3 | 1 575 | 52.32 | 4.67 |
| über 0.3 - 0.4 | 1 698 | 37.46 | 5.67 |
| über 0.4 - 0.5 | 1 179 | 21.29 | 6.88 |
| über 0.5 - 0.6 | 1 188 | 8.92 | 8.32 |
| über 0.6 - 0.7 | 1 505 | 3.39 | 9.77 |
| über 0.7 - 0.8 | 1 524 | 1.05 | 11.12 |
| über 0.8 - 0.9 | 1 658 | 0.00 | 12.68 |
| über 0.9 | 6 271 | 0.03 | 14.67 |
| Total | 27 000 | 76.53 | 2.35 |

^{a)} Der theoretisch erwartete Anteil der Fehler zweiter Art beträgt für $\iota^A = 0.15$ und ein Signifikanzniveau von 95% 0 bis 1.1 Prozent.

realisierte Anteil der Fehler zweiter Art unter einem Prozentpunkt liegt. Theoretisch ist bei dieser Parameterkombination mit einem Anteil bis zu einem Prozent zu rechnen (vgl. Tabellenfussnote oder Tabelle B.1 im Anhang B). Für Werte von $\Delta < 0.7$ liegt der Fehleranteil bereits über dem maximal zu erwartenden Wert und für $\Delta < 0.1$ steigt der Anteil der realisierten Fehler auf über 75 Prozent und übersteigt den erwarteten Anteil von maximal 1.1 Prozent bei Weitem.

Bei Parameterkombinationen mit $\iota^A = 0.20$ wird ein Anteil der Fehler zweiter Art unter einem Promille erwartet (vgl. Tabellenfussnote oder Tabelle B.1 im Anhang B). Für $\Delta > 0.9$ werden bei den Hypothesentests tatsächlich keine Fehler realisiert. Der Anteil der Fehler nimmt wieder mit abnehmenden Werten für Δ kontinuierlich zu; bereits für $\Delta < 0.8$ übersteigt der Anteil der realisierten Fehler zweiter Art den erwarteten Anteil von einem Promille. In der untersten Wertekategorie des Validitätsmasses werden in 59 Prozent aller getesteten Hypothesen Fehler zweiter Art realisiert; wieder ein Wert, der in keinem Verhältnis zum erwarteten Fehleranteil von einem Promille bei der entsprechenden Parameterkombination steht.

In Abschnitt 4.5.1 wurde gezeigt, dass die geschätzten erwarteten Verzerrungen der Effizienzdifferenzen \mathcal{B}^{AB} in einem negativen Zusammenhang mit dem Validitätsmass steht.

Tabelle 4.12: Anteil β -Fehler ($\iota^A = 0.20$)

| Delta Δ | I | Anteil β -Fehler ^{a)} | $\iota^A - \mathcal{B}^{AB}$ |
|----------------|--------|--------------------------------------|------------------------------|
| 0 bis 0.1 | 6 604 | 58.77 | 4.53 |
| über 0.1 - 0.2 | 2 528 | 49.13 | 5.19 |
| über 0.2 - 0.3 | 1 581 | 31.18 | 6.25 |
| über 0.3 - 0.4 | 1 892 | 14.38 | 7.76 |
| über 0.4 - 0.5 | 1 502 | 6.33 | 9.16 |
| über 0.5 - 0.6 | 1 238 | 1.37 | 10.99 |
| über 0.6 - 0.7 | 1 482 | 0.54 | 12.96 |
| über 0.7 - 0.8 | 1 490 | 0.27 | 14.62 |
| über 0.8 - 0.9 | 1 744 | 0.00 | 16.51 |
| über 0.9 | 6 939 | 0.00 | 19.45 |
| Total | 27 000 | 22.67 | 11.10 |

^{a)} Der theoretisch erwartete Anteil der Fehler zweiter Art beträgt für $\iota^A = 0.20$ und ein Signifikanzniveau von 95% 0 bis 1 Promille.

In diesem Abschnitt wurde nachgewiesen, dass sich die verzerrt geschätzten Effizienzdifferenzen auf die Stufe der Hypothesentests übertragen. Der Anteil der Fehler zweiter Art nimmt, unabhängig vom Ausmass der wahren Effizienzunterschiede, mit zunehmenden Werten des Validitätsmasses ab. Diese Resultate stützen die Hypothese 2 (vgl. S. 70).

Folgerung 4.6 *Verzerrt geschätzte Effizienzdifferenzen übertragen sich auf die Ebene der Hypothesentests. Kleine Werte des Validitätsmasses Δ führen zu einer drastischen Zunahme der Fehler zweiter Art. Die in der Simulation realisierten Fehlerhäufigkeiten stimmen gut mit den theoretisch erwarteten Werten für die jeweils realisierten Effizienzdifferenzen überein.*

4.6 Verallgemeinerung der Resultate

In diesem Abschnitt wird der Versuch unternommen, die bisherigen Resultate, die auf einer Ein-Output-Zwei-Input-Transformationsfunktion basieren, auf Szenarien mit weniger restriktiven Annahmen bezüglich der Transformationsfunktion zu übertragen. Die zuverlässigste – aber auch sehr aufwendige – Vorgehensweise ist, entsprechende Simulationen durchzuführen. Darauf wird an dieser Stelle verzichtet, da diese zusätzlichen

Simulationen den Rahmen dieser Arbeit sprengen würden. Stattdessen soll eine qualitative Argumentation die allgemeine Gültigkeit der Folgerungen dieses Kapitels plausibel machen.

Die Möglichkeit, unabhängige Abschnitte im Input-Outputraum zu definieren, nimmt mit der Dimension, d.h. mit der Anzahl Inputs und Outputs, zu. Für eine Transformationsfunktion mit lediglich einem Input und einem Output besteht diese Möglichkeit überhaupt nicht, d.h. Δ ist *immer* gleich eins. Im Ein-Output-Zwei-Input-Fall ist dies – wie oben gezeigt wurde – aber durchaus und in einem nicht zu unterschätzenden Ausmass möglich. Würde nun ein zweiter Output produziert, könnten sich die beiden Gruppen *zusätzlich* auch noch in Bezug auf die Outputkombination unterscheiden, das Validitätsmass Δ kann daher nur abnehmen. Mit jedem weiteren Input oder Output werden zusätzliche Möglichkeiten geschaffen, unabhängige Abschnitte der effizienten Grenze zu definieren.

Folgerung 4.7 *Das Problem der Unabhängigkeit der Abschnitte der GPM in Bezug auf ein Gruppenmerkmal dürfte mit zunehmender Dimension des Input-Outputraumes an Bedeutung gewinnen. Diese Erkenntnis ist umso gravierender, als dass die erwarteten Verzerrungen bei gegebener Stichprobengrösse mit der Dimension zunehmen (da die Konvergenzrate negativ von der Dimension beeinflusst wird, vgl. Gleichung 4.2 auf S. 74).*

Die Möglichkeit, unabhängige Abschnitte zu definieren, ist die Konsequenz der wenig restriktiven Strukturannahmen der DEA. So ist z.B. der Übergang von einem effizienten Abschnitt der GPM auf einen benachbarten Abschnitt in Bezug auf die Grenzzraten der Substitution und Transformation nur durch die Konvexitätsannahme eingeschränkt. Im Gegensatz dazu ist bei den parametrischen Alternativen der Effizienzmessung der Verlauf der Grenze durch den unterstellten funktionalen Zusammenhang stärker vorgegeben. Die Konsequenz ist, dass gruppenspezifische Ineffizienz besser erkannt werden kann und die geschätzten Effizienzdifferenzen eher unverzerrt sind. Es ist zu vermuten, dass das Ausmass der Flexibilität der parametrischen Alternativen entscheidend dafür verantwortlich ist, ob die geschätzten Effizienzdifferenzen verzerrt oder unverzerrt sind. Je flexibler eine parametrische Methode ist, desto grösser dürfte die Gefahr nicht valider Effizienzdifferenzen sein.

Folgerung 4.8 *In vielerlei Hinsichten stellen die wenig restriktiven Strukturannahmen (bzw. die Flexibilität) der DEA einen Vorteil dar. Die Gefahr der verzerrt geschätzten Effizienzdifferenzen stellt aber gewissermassen die Kehrseite dieser Medaille dar.*

Restriktivere Strukturannahmen dürften somit *ceteris paribus* die Gefahr der verzerrt geschätzten Effizienzdifferenzen vermindern. Es ist zu erwarten, dass DEA-Versionen mit konstanten Skalenerträgen weniger anfällig auf verzerrt geschätzte Effizienzdifferenzen sind als solche mit variablen Skalenerträgen (vgl. Abschnitte 2.2.3 und 2.2.3, vgl. aber auch Abschnitt 5.2.1 in Kapitel 5). Auch zusätzlichen Restriktionen auf den virtuellen Multiplikatoren, wie bei der *Cone-Ratio*-DEA, sollten diese Gefahr vermindern (vgl. Abschnitt 2.2.3).

Zum Abschluss dieses Kapitels soll noch die Verbindung zur Kritik am *Two Stage Approach* hergestellt werden (vgl. Abschnitt 2.3.2). Dort wurde darauf hingewiesen, dass eine starke Korrelation zwischen den Inputs bzw. Outputs und den Regressoren eine mögliche Ursache für verzerrt geschätzte Effizienz sei. Als Begründung dafür wird in der Literatur das *Omitted Variable Problem* aufgeführt. Die Erkenntnisse des dritten Kapitels, die hier durch die Simulation bekräftigt werden, zeigen, dass gruppenspezifisch unabhängige Abschnitte der GPM zu verzerrt geschätzten Effizienzdifferenzen führen können. Die Frage stellt sich, ob die Information bezüglich der Validität der Effizienzdifferenzen nicht anhand einer Korrelationsanalyse zwischen den verschiedenen Inputs, Outputs und Regressoren der zweiten Stufe gewonnen werden kann. Eine Begründung, wieso die Korrelationsanalyse in diesem Kontext grundsätzlich nicht geeignet ist, findet sich in Abschnitt 3.4.1. Trotzdem soll an dieser Stelle noch ein zusätzliches Argument für das hier vorgeschlagene Validitätsmass vorgebracht werden. Der Vorteil des Validitätsmasses im Gegensatz zu einer Korrelationsanalyse besteht darin, dass pro Gruppenmerkmal *eine* Kennzahl Auskunft über die Eigenschaften der GPM und die Validität der geschätzten Effizienzdifferenzen gibt. Bei einer Korrelationsanalyse ist für jedes Gruppenmerkmal die Auswertung von $k = \frac{n(n-1)}{2}$ Korrelationskoeffizienten nötig, wobei n die Anzahl der Dimensionen im Input-Outputraum ist.²³ Zudem liesse sich die Frage, ob die verschiedenen Abschnitte der GPM in Bezug auf ein Gruppenmerkmal unabhängig sind, anhand der verschiedenen Korrelationskoeffizienten kaum beantworten.

Folgerung 4.9 *Das Validitätsmass Δ ist verhältnismässig einfach zu berechnen und liefert pro Gruppenmerkmal zudem eine Kennzahl zur Bewertung der Validität der Effizienzdifferenzen.*

²³Werden die Input-, Output- und Input-Outputverhältnisse ausgewertet, diese sind für die implizite Konsistenzbedingung relevant, müssen gar $K = \frac{k(k-1)}{2}$ Korrelationskoeffizienten ausgewertet werden, wobei wiederum $k = \frac{n(n-1)}{2}$ ist. Dies erschwert die Beantwortung der Frage, da $K \gg k$.

Kapitel 5

Eine empirische Anwendung der DEA im schweizerischen Krankenhaussektor

In diesem Kapitel werden die bisherigen Erkenntnisse anhand einer empirischen Arbeit illustriert. Im ersten Abschnitt wird eine Anwendung der DEA zur Effizienz der schweizerischen Krankenhäuser präsentiert. Im zweiten Abschnitt folgt eine Analyse der Resultate der ökonometrischen Auswertung der DEA-Effizienz. Es wird insbesondere untersucht, ob Hinweise für verzerrt geschätzte Effizienzdifferenzen vorliegen, die zu einer erhöhten Wahrscheinlichkeit von Fehlern zweiter Art führen. Dabei wird den Erkenntnissen der Kapitel 3 und 4 besondere Aufmerksamkeit gewidmet.

5.1 DEA schweizerischer Krankenhäuser

Die präsentierte Studie basiert auf einer empirischen Anwendung der DEA im schweizerischen Krankenhaussektor von Steinmann und Zweifel (2000). Im Gegensatz zur ursprünglichen Studie wird sie hier erweitert: Anstelle der traditionellen DEA wird eine Versionen mit restringierten Grenzzahlen der Substitution und Transformation zur Anwendung kommen (die sogenannte *Cone-Ratio*-DEA, vgl. Abschnitt 2.2.3).

5.1.1 Die Inputs und Outputs der Krankenhäuser

Als erstes muss geklärt werden, mit welchen Inputs im Krankenhaussektor welche Outputs produziert werden. Einigkeit besteht grösstenteils bei den Inputs, die im wesentlichen Arbeit und Kapital umfassen. Das zeigt sich in den bisherigen Anwendungen der DEA auf Krankenhäuser. Die Auswahl der Inputs stimmt weitgehend überein, wobei bis zu fünf Personalkategorien unterschieden werden.

In Ermangelung zuverlässiger buchhalterischer Daten, die eine zuverlässige Bestimmung des Kapitaleinsatzes ermöglichen, dient üblicherweise die Zahl der Betten eines Krankenhauses als Indikator für das Kapital. Burgess und Wilson (1996) nehmen eine Unterteilung in Langzeit- und Kurzzeitbetten vor, um der unterschiedlichen Ausstattung mit medizinischer Technologie und damit dem Kapital Rechnung zu tragen. Byrnes und Valdmanis (1994) argumentieren dagegen, dass für den Transformationsprozess eher das effektiv eingesetzte Kapital relevant ist, in diesem Kontext wären dies die effektiv belegten Betten. Allerdings wird so eine Quelle der Ineffizienz nicht erfasst, nämlich die Nichtnutzung von vorhandenem Kapital. Ein entsprechender Überschuss lässt sich aber auch als Vorhaltung von Betten und somit als Output interpretieren [vgl. Zweifel und Breyer (1997), Kap. 9]. Aus dieser Sicht erscheint die Arbeit von Banker, Conrad und Strauss (1986) in einem günstigen Licht, weil dort die Kapitalnutzungskosten, d.h. die Summe der Verzinsung und der Abschreibungen, selbst als Input des Krankenhauses verwendet wird. Allerdings ist zu beachten, dass die Kapitalnutzungskosten öffentlicher Unternehmen i.d.R. keinen eindeutigen Rückschluss auf das eingesetzte Kapital zulässt. Färe, Grosskopf, Lindgren und Roos (1994) verzichten mangels verfügbarer Informationen gänzlich auf die Verwendung von Kapital als Produktionsfaktor und führen neben der Arbeit den Sachaufwand als weiteren Input ein.

Auf der Seite der Outputs präsentiert sich die Sachlage nicht gleichermassen eindeutig. Unumstritten scheint, dass die Leistung der Krankenhäuser darin besteht, die Gesundheitsversorgung im weiteren Sinn zu gewährleisten. Was unter dem Begriff der „Gesundheitsversorgung im weiteren Sinn“ subsumiert wird ist aber kontrovers. Die Ausbildung von Personal oder die medizinische Grundlagenforschung können, müssen aber nicht, in diesen Leistungskatalog fallen. Ein weiteres Problem besteht insbesondere bei der Operationalisierung der Outputs: Da der wahre Beitrag der Krankenhäuser zum Gesundheitszustand der Patienten – die Differenz zwischen dem Gesundheitszustand nach der Behandlung und demjenigen, der sich ohne Behandlung eingestellt hätte – überhaupt

nicht gemessen werden kann, muss zwangsläufig auf Outputindikatoren zurückgegriffen werden.

Bei allen bisherigen Studien fehlen Indikatoren des Beitrags der Krankenhäuser zur Gesundheit der Patienten. Üblicherweise behilft man sich mit jenen Outputdimensionen, die Gegenstand der Honorierung sind, also Pflegetage und Behandlungsfälle. Eine Maximierung der Pflegetage kann aus volkswirtschaftlicher Sicht jedoch kaum als Zeichen einer effizienten Betriebsführung aufgefasst werden, würde doch damit ein wichtiger Input in Form der Zeit der Patienten verschwendet. Hingegen macht eine Unterteilung der Behandlungsfälle in verschiedene Kategorien Sinn, beispielsweise in ambulanten Untersuchungen und Behandlungen oder auch in ambulante, stationäre und gynäkologische Fälle. Eine zu weit gehende Unterteilung der Inputs und Outputs hätte allerdings zur Konsequenz, dass die Zahl der Abschnitte der GPM zunimmt, so dass ein grösserer Anteil der Beobachtungen als effizient erkannt wird.¹

Burgess und Wilson (1996) unterscheiden zwischen Pflegetagen, Langzeitpflegetagen sowie ambulanten und stationären Behandlungen und Entlassungen (z.T. mit dem *Case Mix* gewichtet). Byrnes und Valdmanis (1994) verwenden dagegen akute, stationäre und gynäkologische Fälle, während Banker et al. (1986) drei vom Alter der Patienten abhängige Pflegetagkategorien favorisieren. In Färe et al. (1994) werden die Outputs in Behandlungsfälle, Untersuchungen (Einzelleistungen) sowie Pflegetage unterteilt. Einzig Linna und Hakkinen (1998) berücksichtigen neben den üblichen Outputs die Forschung sowie die Ausbildung als zusätzliche Leistungen der Krankenhäuser.

In der hier präsentierten DEA [in Analogie zu Steinmann und Zweifel (2000)] werden die Beschäftigten dreier Personalkategorien (akademisches Personal, Pflegepersonal sowie Personal der administrativen und technischen Dienste) als Inputs unterschieden. Als weiterer Input dient der Sachaufwand (zu Preisen von 1990). Da die Krankenhäuser für Medikamente und Hilfsmittel landesweit die gleichen Marktpreise bezahlen, ist der Rückschluss von dieser Wertgrösse auf die für die Produktion verwendeten Mengen durchaus vertretbar. In Übereinstimmung mit obiger Argumentation figurieren schliesslich die Pflegetage unter den Inputs. Auf der Seite der Outputs werden nur stationäre Leistungen erfasst, obschon insbesondere Krankenhäuser in Städten durch ihre Ambulatorien auch ambulante Leistungen erbringen. Die Behandlungsfälle dürfen als einigermaßen homogen

¹Werden zusätzliche Variablen in ein zu maximierendes Lineares Programm eingeführt, nehmen die optimalen Lösungen mindestens gleich grosse Werte an, d.h. im Durchschnitt nimmt der Wert der Lösungen zu.

gelten, werden doch die folgenden fünf Kategorien unterschieden: (1) medizinische, (2) pädiatrische, (3) chirurgische, (4) der Intensivpflege zugeordnete und (5) gynäkologische Behandlungsfälle.

5.1.2 Die Daten der Stichprobe

Für den Zeitraum von 1993-96 liegen 750 Beobachtungen von insgesamt 190 Allgemeinkrankenhäusern vor, welche im Rahmen der administrativen Statistik von der *H+ die Spitäler der Schweiz* erhoben wurden.

Fehlerhafte Daten wurden im Rahmen der Datenbereinigung aus der Stichprobe entfernt. Bei der Kontrolle des Datensatzes ergaben sich bei Krankenhäusern mit Langzeitpatienten z.T. mehrere Tausend verrechnete Pflgetage pro Patient und Jahr. Da die Behandlungsfälle als Entlassungen definiert sind, müssten Langzeitpatienten – die bei den psychiatrischen oder geriatrischen Patienten auftreten – am Ende jedes Jahres formell entlassen werden und im neuen Jahr formell wieder eintreten. Krankenhäuser verzichten aber oft auf diese umständliche Prozedur und weisen in der Folge systematisch zu wenig Behandlungsfälle aus. Da nicht bekannt ist, welche Krankenhäuser systematisch zu wenig Behandlungsfälle bei den Langzeitpatienten aufweisen, wurden alle Krankenhäuser, die Langzeitpatienten behandeln, von der Studie ausgeschlossen. Somit bleiben von ursprünglich 750 Beobachtungen von 190 Krankenhäusern über die Periode 1993-96 noch 310 Beobachtungen von 89 Krankenhäusern übrig.

5.1.3 Die Restriktionen der Grenzzraten der Substitution und Transformation

Um zu verhindern, dass bei der Lösung der Zielfunktion der DEA die virtuellen Multiplikatoren unplausible Werte annehmen, werden – wie in Abschnitt 2.2.3 erläutert – die zulässigen Werte der Grenzzraten der Substitution und Transformation beschränkt. Die unterstellten minimalen und maximalen Werte für die fünf Inputs sind in der Tabelle 5.1 aufgeführt. Diese Schranken mögen auf den ersten Blick restriktiv wirken. Es ist aber zu beachten, dass es sich dabei um Durchschnittswerte pro Beobachtung und Jahr handelt.²

²Alternative Restriktionen auf den Grenzzraten der Substitution und Transformation wurden getestet. Dabei hat sich gezeigt, dass die geschätzte Effizienz nur schwach auf eine Veränderung der Restriktionen reagiert.

Tabelle 5.1: Unterstellte Wertgrenzen der Produktionsfaktoren, Fr. zu Preisen 1990

| | minimaler Wert | maximaler Wert |
|---------------------------|----------------|----------------|
| Akademiker | 75 000 | 350 000 |
| Pflegepersonal | 30 000 | 120 000 |
| Verwaltungspersonal | 30 000 | 120 000 |
| Pflegetage | 100 | 2000 |
| Sachaufwand ^{a)} | 1 | 1 |

^{a)} Der Sachaufwand dient als *Numéraire* und ist auf den Wert eins fixiert.

Zum Vergleich: Das durchschnittliche Jahressalär der Angestellten im Gesundheitssektor der Schweiz belief sich im Jahr 1994 auf Fr. 48 400.- [vgl. OECD (1999)]. Das durchschnittliche Einkommen der Ärzte im selben Jahr betrug Fr. 146 000.- (jeweils zu Preisen von 1990). Da die überwiegende Mehrheit der Angestellten im Pflegebereich tätig ist, ist die Lohnverteilung im Gesundheitssektor linkssteil, m.a.W. der Durchschnittslohn in den Pflegeberufen muss sogar unter 48 400.- liegen.³

Da die Pflegetage als Input seitens der Patienten definiert wurden, dienen die Opportunitätskosten eines Pflegetages als Schranken für die Grenzrate der Substitution. Die Opportunitätskosten fallen an, da der Patient während der stationären Behandlung nicht seiner beruflichen, gesellschaftlichen oder familiären Tätigkeit nachgehen kann. Werden die unterstellten Werte pro Pflegetag auf Monatswerte hochgerechnet (vgl. vierte Zeile der Tabelle 5.1), ergibt sich ein minimaler Wert von rund Fr. 3 000.- bzw. ein maximaler Wert von rund Fr. 600 000. Die untere Begrenzung entspricht somit in etwa dem Existenzminimum, die obere Schranke repräsentiert dagegen das oberste Perzentil der Lohnverteilung.

Die virtuellen Multiplikatoren der Outputs und somit die Grenzzraten der Transformation wurden ebenfalls restringiert. Begründet wird diese Massnahme damit, dass bei einer Spezialisierung auf einen oder wenige Outputs zur Erreichung einer vorgegebenen Effizienz auch eine Zunahme der entsprechenden Outputs notwendig sein soll. Ein Beispiel zur Illustration: Verglichen werden zwei Krankenhäuser mit identischen Outputsets (vgl. Abschnitt 2.1). Das eine Haus produziert jeweils 1000 medizinische und gynäkologische Fälle. Das zweite Haus hat sich auf medizinische Fälle spezialisiert, d.h. es produziert ausschliesslich medizinische Fälle. Ohne Restriktionen auf den Grenzzraten der Transfor-

³Im Jahr 1987 waren von den 66 650 im Gesundheitssektor angestellten Personen nur gerade 9 950 praktizierende Ärzte.

mation genügten dem zweiten Haus 1001 medizinische Fälle, um effizient nach Pareto-Koopmans zu sein. Der marginale medizinische Fall wäre in dieser Situation gleich viel wert wie 1000 gynäkologische Fälle. In der vorliegenden Studie muss der durchschnittliche Wert pro Behandlungsfall einer Kategorie mindestens einem Zehntel des Wertes pro Behandlungsfall der übrigen Kategorien entsprechen. Im Beispiel bedeutet dies, dass das spezialisierte Haus mindestens 1100 medizinische Fälle produzieren muss, damit es als effizient deklariert wird.

Da die Wertebereiche aller virtuellen Multiplikatoren auf positive Werte beschränkt werden, ist das Effizienzmass strikt monoton und erlaubt die Effizienzindikation (vgl. Abschnitt 2.2) und die GPM besteht ausschliesslich aus pareto-koopmans-effizienten Abschnitten.

5.1.4 Die Ergebnisse der DEA

Die Ergebnisse der DEA werden nicht explizit aufgeführt, da diese ohne Berücksichtigung der Rahmenbedingungen wenig aussagekräftig sind.⁴ Erst die Berücksichtigung der heterogenen Rahmenbedingungen ermöglicht volkswirtschaftlich relevante Aussagen. Zu erwähnen ist einzig, dass Beobachtungen aller Grössenklassen effiziente Beobachtungen stellen. Die einzige Ausnahme diesbezüglich stellen die 33 Beobachtungen der Krankenhauskategorie mit 250-499 Betten dar. Dies kann als Rechtfertigung für die Strukturannahme der konstanten Skalenerträge angesehen werden. Trotz dieser Restriktion sind Beobachtungen der kleinsten Krankenhauskategorie (weniger als 75 Betten) wie auch Universitätsspitäler in der Lage, effiziente Abschnitte der GPM zu definieren. Da in einer mittleren Krankenhauskategorie keine Beobachtungen als effizient eingestuft werden, stellt sich die Frage, ob der schweizerische Krankenhaussektor im beobachteten Input-Outputbereich durch positive Skalanelastizität, d.h. global zunehmende Skalenerträge, gekennzeichnet ist.⁵

⁴Die DEA wurde mit *Efficiency Measurement System* (EMS, Release1.2) durchgeführt. Ich bedanke mich bei Herrn Dr. H. Scheel, dem Entwickler der Software, der immer umgehend und kompetent meine Fragen zu EMS beantwortet hat.

⁵Ist die zugrunde liegende wahre Technologie durch global positive Skalanelastizität gekennzeichnet, kann die Effizienz anhand der DEA nicht konsistent geschätzt werden, da die Strukturannahme der konvexen Technologiemenge der DEA nicht gerechtfertigt ist [vgl. Simar und Wilson (2000)]. Die Ergebnisse wären dann im mittleren Grössenbereich nach unten, d.h. in Richtung zu ineffizient, verzerrt.

5.1.5 Die ökonometrische Auswertung der DEA-Ergebnisse

Wie in Kapitel 2 erwähnt wurde, ist es kaum möglich, alle die Input-Outputtransformationen beeinflussenden Rahmenbedingungen in die Schätzung der Effizienz zu integrieren. In diesem Abschnitt wird deshalb der Versuch unternommen, die relative Distanz zur pareto-koopmans-effizienten GPM durch ein Set von Variablen, das die Rahmenbedingungen erfassen soll, zu erklären. Insbesondere stellt sich die Frage, wie sich die Anreizstrukturen im schweizerischen Krankenhaussektor auf die Effizienz der Leistungserbringung auswirken. Folgende ökonomischen Hypothesen sollen überprüft werden:

- *Defizitdeckung.* Falls die Defizite sanktionslos gedeckt werden ($DEFD=1$), entfallen die Anreize zur effizienten Produktion. Die Defizitdeckung ist mit geringerer Effizienz in Verbindung zu bringen.
- *Subventionierung.* Falls eine Beobachtung subventioniert wird ($SUBV=1$) und das Budget festgelegt ist, bestehen keine Anreize für effizientes Wirtschaften. Es könnte sogar im Interesse der Entscheidungsträger sein, Ressourcen zur Rechtfertigung einer Budgetausweitung für die Folgeperioden verschwenderisch einzusetzen. Die Subventionierung ist daher mit geringerer Effizienz in Verbindung zu bringen. Zu beachten ist zudem, dass eine Subventionierung eine komplementäre Defizitdeckung nicht zwingend ausschliesst.
- *Subventionierung ohne komplementäre Defizitdeckung.* Falls eine Beobachtung subventioniert wird und die nachträgliche Deckung eines allfälligen Defizits ausgeschlossen ist ($SUBV(1-DEFD)=1$), bestehen Anreize zur effizienten Produktion. Dies gilt insbesondere dann, wenn allfällige Überschüsse von den Entscheidungsträgern zweckgebunden, z.B. für Investitionen, eingesetzt werden können.
- *Rechtsträger.* Die Hypothese besagt, dass die Entscheidungsträger der privatrechtlich organisierten Beobachtungen ($OEFF=0$) besser, d.h. mit geringeren Kosten von den Eigentümern kontrolliert und gegebenenfalls sanktioniert werden können. Daher bestehen für die Entscheidungsträger privatrechtlich organisierter Beobachtungen stärkere Anreize zur effizienten Leistungserbringung als für Entscheidungsträger öffentlich-rechtlich organisierter Beobachtungen ($OEFF=1$). Zu beachten ist jedoch, dass auch Beobachtungen mit privatrechtlicher Rechtsform Eigentum der öffentlichen Hand sein können.

- *Privatspital*. $PRIV=1$, falls die Beobachtungen sich als Privatspital deklariert und dem *Schweizerischen Verein Privater Spitäler* (SVPK) angehört. Werden gewisse Auflagen erfüllt, i.d.R. muss eine Abteilung für allgemeinversicherte Patienten betrieben werden, können die Privatspitäler Subventionen von der öffentlichen Hand bekommen. Der erwartete Zusammenhang mit der Effizienz ist unbestimmt.

Weitere statistische Kontrollvariablen wurden zur Erfassung der unterschiedlichen Rahmenbedingungen in die Schätzung eingeführt:

- *Anteil der Angestellten in Ausbildung*. Ein grosser Anteil Auszubildender (AUSB) sollte tendenziell die Effizienz senken. Dieses Postulat basiert auf der Annahme, dass Auszubildende vergleichsweise weniger produktiv sind. Trotzdem werden sie in der jeweiligen Personalkategorie ohne Einschränkungen zum Bestand gezählt. Zudem besteht in der hier präsentierten DEA nicht die Möglichkeit, die niedrigere Produktivität mit geringeren Personalkosten zu kompensieren. Um nicht-lineare Zusammenhänge feststellen zu können, wird zudem der quadrierte Anteil in die Schätzgleichung eingeführt (AUSB2).
- *Anteil der Pfl egetage der nicht-allgemeinen Abteilungen am Total der verrechneten Pfl egetage*. Diese Variable ($PRIV\%$) soll Qualitätsunterschiede in der Leistungserbringung auffangen, die kaum erfasst werden können und in der hier vorliegenden DEA unberücksichtigt bleiben. Geht man davon aus, dass Qualität zusätzliche Ressourcen verzehrt, wird eine höhere Qualität von der DEA als Ineffizienz interpretiert. Steht den Patienten z.B. in einem Krankenhaus permanent ein Heer von Pflegepersonal bereit, wird dies von den Patienten vermutlich als Qualität aufgefasst. Dasselbe dürfte auch auf den Input Kapital zutreffen: Ein Krankenhaus, das ausschliesslich Einzelzimmer anbietet, wird vermutlich eine niedrigere Bettendichte aufweisen und folglich weniger Behandlungsfälle pro Kapitaleinheit hervorbringen, was von der DEA als Ineffizienz erkannt würde.
- *Die Patientenimport-Exportquote*. IMEXQ: Das Verhältnis zwischen der Anzahl Patienten, die von anderen Kantonen überwiesen wurden, zur Anzahl Patienten, die in Krankenhäuser anderer Kantone überwiesen wurden. Zwei Gründe sprechen für einen negativen Zusammenhang mit der DEA-Effizienz. Zum einen müssen Kantone, deren Krankenhauskapazität ausgelastet ist, Patienten exportieren. Eine hohe

Auslastung der Kapazität erhöht *ceteris paribus* die Effizienz. Zum anderen exportieren Kantone, die keine Spitzenmedizin anbieten, ihre auf Spitzenmedizin angewiesenen Patienten in Kantone, die diese auch anbieten. Da die Spitzenmedizin pro Behandlungsfall besonders ressourcenintensiv ist, spricht auch dieses zweite Argument für einen negativen Zusammenhang mit der DEA-Effizienz. Die Daten zur Patientenimport-Exportquote entstammen Crivelli (1998).

- *Notfallaufnahme*. NOTFALL=1, falls die Beobachtung über eine Notfallaufnahme verfügt. Eine Notfallaufnahme kann sich auf zwei verschiedene Arten auf die DEA-Effizienz auswirken. Einerseits wird die Planung der Leistungserbringung durch das stochastische Eintreffen von Notfällen zu einem gewissen Grad erschwert. Die Folge ist, dass zusätzliche Kapazitäten für Notfälle freigehalten werden müssen und die Auslastung insgesamt abnimmt. Somit sollte die DEA-Effizienz bei Beobachtungen mit Notfallaufnahme *ceteris paribus* kleiner ausfallen. Andererseits ist auch denkbar, dass einer Beobachtung über die Notfallaufnahme „günstige“, d.h. einfache Fälle zukommen, die die DEA-Effizienz steigern.
- *Jahr der Beobachtung*. JAHRJJ: Diese Dummyvariablen nehmen den Wert eins an, wenn die Beobachtung ins Jahr JJ fällt. Der Beobachtungszeitraum war geprägt von einer Revision des Krankenversicherungsgesetzes, das auf Anfang 1995 eingeführt wurde, und einem aufkommenden und in den Medien breit diskutierten Kostenbewusstsein. Um die zeitliche Entwicklung der DEA-Effizienz zu erfassen, wurden Dummyvariablen für die Beobachtungsjahre 1994 bis 96 in die Schätzgleichung eingeführt (als Basisjahr dient 1993).
- *Krankenhauskategorie*. KATg: Es wurde eine Technologiemenge mit konstanten Skalenerträgen unterstellt. Mit Dummyvariablen, die die Grösse der Beobachtung erfassen, soll untersucht werden, ob die DEA-Effizienz mit der Krankenhausgrösse zusammenhängt. Wird ein systematischer Zusammenhang festgestellt, muss die Strukturannahme der konstanten Skalenerträge in Frage gestellt werden. Die Dummyvariablen nehmen den Wert eins an, falls die Beobachtung in die Krankenhauskategorie g fällt; die Universitätsspitäler bilden die Basiskategorie g=1. Zusätzlich gelten g=2: 500 und mehr Betten, g=3: 250-499 Betten, g=4: 125-249 Betten, g=5: 75-124 Betten und g=6: weniger als 75 Betten.

Bei der Erklärung der DEA-Effizienz gilt es, zwei Eigenheiten zu beachten. Erstens liegen Paneldaten vor,⁶ was zur Folge hat, dass die Annahme, wonach die Beobachtungen der Stichprobe unabhängig sein müssen, verletzt wird. Aus diesem Grund wird bei den Panel-Schätzmethoden eine individuumspezifische Variable in die Schätzgleichung eingeführt. Unbeobachtete individuumspezifische Heterogenität wird somit kontrolliert und es wird verhindert, dass sie den systematischen Zusammenhängen zugeordnet wird. Grundsätzlich bestehen zwei Methoden, individuumspezifische Variablen in eine Schätzgleichung zu integrieren. Die erste besteht darin, für jedes Individuum eine Dummyvariable einzuführen (zur Vermeidung von perfekter Kollinearität wird gleichzeitig die Konstante aus der Schätzgleichung entfernt). Dieses Verfahren wird als *Fixed Effects Regression* bezeichnet. Die gesamte unbeobachtete individuumspezifische Heterogenität wird dann durch diese Dummyvariablen aufgefangen. Der Vorteil der *Fixed Effects Regression* besteht darin, dass keine Annahmen bezüglich der Verteilung der geschätzten Dummyvariablen nötig sind. Ein Nachteil ist allerdings, dass pro Dummyvariable und somit pro Individuum ein Freiheitsgrad geopfert werden muss, was gerade bei kleinen Stichproben einen hohen Preis darstellt. Zudem kann nur die Wirkung von erklärenden Variablen geschätzt werden, die über die Zeit variieren, da sonst perfekte Kollinearität mit den individuumspezifischen Konstanten besteht. Bei der zweiten Methode, der sogenannten *Random Effects Regression*, wird die Heterogenität durch einen zusätzlichen, individuumspezifischen Störterm mit einem Erwartungswert von null erfasst. Der Nachteil dieses Verfahrens besteht darin, dass unterstellt wird, der individuumspezifische Störterm sei mit den Regressoren unkorreliert.⁷

Weiter ist zu beachten, dass die Verteilung der DEA-Effizienz typischerweise nicht den Standardannahmen der Regression entspricht. Da die Verteilung der Effizienz bei 100% abgeschnitten ist, resultiert ein Selektionseffekt. Die Stichprobe lässt sich in vollständig effiziente und ineffiziente Beobachtungen unterteilen, wobei ausschliesslich bei der zweiten Gruppe das Ausmass der Effizienz variiert. Wird in einer solchen Situation die abhängige Variable (hier die Effizienz) ohne entsprechende Korrektur geschätzt, resultiert ein Störterm mit einem Erwartungswert ungleich null, was zu inkonsistenten und verzerrt geschätzten Koeffizienten führt. Eine Schätzmethode, die diese Verzerrung korrigiert, ist das *Tobit*-Verfahren [vgl. Greene (1997), Kap. 20].

⁶Ein Paneldatensatz ist eine kombinierte Querschnitts-Längsschnittstichprobe. Dies ist z.B. dann der Fall, wenn pro Individuum – bzw. pro Krankenhaus – in der Stichprobe mehrere Beobachtungen vorliegen.

⁷Eine Einführung in die Panelschätzmethoden findet sich in Greene (1997), Kap. 14 oder Stata (2001), *Reference Su-Z*, S. 434ff.

Die Auswertung der DEA-Effizienz erfordert somit eine Schätzmethode, die sowohl die Panelstruktur der Daten als auch die abgeschnittene Verteilung der abhängigen Variable (Effizienz) berücksichtigt. Die ökonometrische Auswertung der DEA-Effizienz anhand des *Random Effects Tobit* hat sich aber bei den vorliegenden Daten als derart unrobust erwiesen, dass auf eine Präsentation und Diskussion der Resultate verzichtet wird.⁸ Das unkonditionierte *Fixed Effects Tobit* – bei dem die *Fixed Effects* als Dummyvariablen in die Matrix der erklärenden Variablen eingehen und dann ein normales *Tobit*-Modell geschätzt wird – führt zu verzerrt geschätzten Koeffizienten [vgl. Stata (2001), *Reference Su-Z*, S. 474]. Zudem konvergierte die *Log-Likelihood*-Funktion bei den vorliegenden Daten nicht. Als Alternative zu den *Tobit*-Verfahren wird daher ein *Two-Part*-Modell geschätzt. Beim *Two-Part*-Modell wird zuerst die Zugehörigkeit zur Gruppe der effizienten Beobachtungen untersucht (mit einem *Random Effects Probit*), dann wird bei der Teilstichprobe der ineffizienten Beobachtungen der Grad der Effizienz analysiert. Die Zugehörigkeit zur Gruppe der effizienten Beobachtungen kann mit den vorhandenen Daten nicht erklärt werden, da das *Random Effects Probit* denselben Approximationsalgorithmus wie das *Random-Effects-Tobit*-Verfahren verwendet und die Resultate wieder sensitiv in Bezug auf die Anzahl Punkte der *Gauss-Hermite*-Funktion sind. Die Analyse der DEA-Effizienz wird deshalb auf die Teilstichprobe der ineffizienten Beobachtungen beschränkt.

In Tabelle 5.2 sind die Resultate der linearen *Random-Effects*-Schätzung der Effizienz der Teilstichprobe der ineffizienten Beobachtungen aufgeführt. Die Teststatistik zeigt, dass die Schätzgleichung insgesamt einen hochsignifikanten Erklärungsgehalt hat. Allerdings fällt das korrigierte Bestimmtheitsmass mit 7.5% eher niedrig aus; die Regressoren vermögen lediglich 7.5% der gesamten Varianz zu erklären. Die geschätzte Varianz des *Random Effects*, der unbeobachteten Heterogenität, macht zudem rund 84% der gesamten Varianz aus. Dies ist ein Indiz, das für die Verwendung der Panelschätzmethode spricht. Der Hausman-Spezifikationstest zeigt, dass weder die *Random Effects* mit den erklärenden Variablen korreliert sind noch eine Missspezifikation vermutet werden muss.⁹ Die beiden

⁸Die Dichtefunktion der bedingten Wahrscheinlichkeitsverteilung der abhängigen Variable wird durch eine *Gauss-Hermite*-Funktion mit m Punkten approximiert. Variieren die Schätzergebnisse bei einer alternativen Anzahl Punkte (m) der *Gauss-Hermite*-Funktion, müssen die Resultate als nicht-robust betrachtet werden. Als Faustregel gilt, dass die Abweichungen kleiner als ein Prozent sein sollten [vgl. Stata (2001), *Reference Su-Z*, S. 420ff. bzw. 477ff.].

⁹Der Hausman-Test baut auf der Annahme auf, dass alle unbeobachteten Effekte bei der *Fixed Effects Regression* im *Fixed Effect* enthalten sind und folglich die geschätzten Koeffizienten unverzerrt sind. Mit dem Hausman-Test wird überprüft, ob sich die Koeffizienten der *Random Effects Regression* signifikant

Tabelle 5.2: *Random-Effects*-Schätzung^{a)} der DEA-Effizienz

| Teilstichprobe der ineffizienten Beobachtungen (n=273) | | | | |
|--|-----------------|---------------------------|--------|-------|
| Variable | erw. Vorzeichen | Koeffizient ^{b)} | z-Wert | P>z |
| DEFD | - | -3.276 | -1.85 | 0.064 |
| SUBV | - | -3.590 | -1.73 | 0.083 |
| SUBV(1-DEFD) | + | 1.147 | 0.39 | 0.699 |
| OEFF | - | -2.626 | -0.87 | 0.385 |
| PRIV | ? | 1.767 | 0.39 | 0.698 |
| AUSB | - | 0.432 | 1.67 | 0.096 |
| AUSB2 | - | -0.012 | -1.09 | 0.275 |
| PRIV% | - | -0.099* | -2.01 | 0.044 |
| IMEXQ | - | -1.529* | -1.98 | 0.048 |
| NOTFALL | ? | 0.411 | 0.09 | 0.932 |
| JAHR94 | ? | 2.189* | 2.46 | 0.014 |
| JAHR95 | ? | 3.271*** | 3.67 | 0.000 |
| JAHR96 | ? | 4.393*** | 4.68 | 0.000 |
| KAT2 | ? | 8.292 | 0.82 | 0.412 |
| KAT3 | ? | 7.729 | 0.92 | 0.358 |
| KAT4 | ? | 9.439 | 1.16 | 0.246 |
| KAT5 | ? | 7.985 | 0.98 | 0.328 |
| KAT6 | ? | 6.263 | 0.76 | 0.448 |
| Konstante | ? | 69.710*** | 6.88 | 0.000 |
| σ_u | 10.137 | | | |
| σ_e | 4.493 | | | |
| $\rho = \frac{\sigma_u}{\sigma_u + \sigma_e}$ | 0.836 | | | |
| adj. R^2 | 0.075 | | | |
| $\chi^2(18)$ | 66.57 | | | |
| $P > \chi^2$ | 0.000 | | | |

^{a)} Die *Generalized Least Squares Random-Effects*-Schätzung erfolgten mit Stata (Release 7.0)

^{b)} *, **, ***: signifikant verschieden von null mit der Wahrscheinlichkeit 95, 99, 99.9%.

von den (unverzerrten) Koeffizienten der *Fixed Effects Regression* unterscheiden. Ist die nicht der Fall, wird vermutet, dass die *Random Effects* nicht mit den erklärenden Variablen korreliert sind und keine

Variablen DEFD und SUBV weisen das erwartete Vorzeichen auf, sind aber knapp nicht signifikant von null verschieden. Beobachtungen mit Defizitdeckung sind gemäss Schätzung im Durchschnitt rund 3 Prozentpunkte weniger effizient als die übrigen Beobachtungen. Noch einen grösseren Effekt scheinen die Subventionen zu haben, die eine um rund 3.5 Prozentpunkte geringere Effizienz nach sich ziehen. Die übrigen die Anreizstruktur erfassenden Variablen (SUBV(1-DEFD), OEFF und PRIV) sind statistisch eindeutig nicht von null zu unterscheiden (vgl. dritte bis fünfte Zeile der Tabelle 5.2). Immerhin scheinen die Pauschalsubventionen [SUBV(1-DEFD)] und die Rechtsform (OEFF) das erwartete Vorzeichen zu haben.

Die Kontrollvariablen AUSB und AUSB2 sind beide nicht signifikant von null verschieden. Ein grösserer Anteil Auszubildender scheint einen positiven Einfluss auf die Effizienz zu haben, der ab einem Anteil von rund 18% aber zu einer Abnahme führt.¹⁰ Der Anteil der Pflgetage der nicht-allgemeinen Abteilungen am Total der Pflgetage (PRIV%) hat wie erwartet einen negativen und zudem signifikanten Einfluss auf die Effizienz. Die Import-Exportquote hat das erwartete Vorzeichen und der Zusammenhang ist ebenfalls statistisch signifikant. Beobachtungen aus Kantonen, die viele Patienten in andere Kantone exportieren, aber selbst wenig importieren, scheinen effizienter zu sein. Schliesslich unterscheiden sich Beobachtungen mit einer Notfallaufnahme in Bezug auf die Effizienz kaum von den übrigen Beobachtungen. Erfreulich ist, dass die Beobachtungen in dieser Stichprobe im Laufe der vier Beobachtungsjahre stetig effizienter wurden. Im Vergleich zum Basisjahr 1993 hat die durchschnittliche Effizienz im Jahr 94 um rund 2.2, im Jahr 95 um rund 3.3 und im letzten Beobachtungsjahr gar um rund 4.4 Prozentpunkte zugenommen. Es ist zu vermuten, dass diese Effizienzsteigerung entweder auf die Revision des KVG zurückzuführen ist oder die Konsequenz einer Produktivitätssteigerung im Krankenhaussektor darstellt. Alle Dummyvariablen, die die Krankenhausgrösse erfassen, sind nicht signifikant von null verschieden. Dies bestätigt die Erkenntnisse des Abschnitts 5.1.4: Die bei der DEA unterstellte Technologie mit konstanten Skalenerträgen scheint auch diesem Test zu genügen.

Folgerung 5.1 *Zusammenfassend muss festgehalten werden, dass die Anreizstrukturen die Varianz in der Effizienz nur unbefriedigend zu erklären vermögen. Einerseits fällt*

Missspezifikation vorliegt.

¹⁰Auf die Interpretation dieser Koeffizienten wird nicht verzichtet, da die mangelnde Signifikanz auf die hohe Korrelation zwischen den beiden Variablen AUSB und AUSB2 zurückzuführen ist. Werden diese Variablen einzeln in die Schätzgleichung integriert, resultieren signifikante Koeffizienten.

das Bestimmtheitsmass mit rund 7.5% eher gering aus. Des Weiteren ist der Anteil der unbeobachteten heterogenen Varianz (ρ) mit rund 85% sehr gross. Zudem sind die statistischen Zusammenhänge der erklärenden Variablen nicht bzw. nur knapp signifikant von null verschieden.

5.1.6 Bewertung der Schätz-Resultate

Mögliche Ursachen für den bescheidenen Erklärungsgehalt der ökonometrischen Auswertung der DEA-Effizienz bieten folgende Interpretationen:¹¹

1. Die Resultate sind ein Indiz dafür, dass die Effizienzunterschiede im schweizerischen Spitalsektor nicht auf die Anreizstrukturen zurückzuführen sind. Die Anreizstrukturen sind dermassen schwach, dass keine Gruppe gezwungen bzw. veranlasst wird, die Leistungserbringung effizienter zu gestalten. Eine mögliche Erklärung dafür wäre, dass Beobachtungen, die effizienter sein könnten, sich an den ineffizienten orientieren und diese imitieren [vgl. Laffont und Tirole (1993), Kap. 2]. Entscheidend wäre in dieser Situation, dass die Hypothesen zu Recht nicht abgewiesen werden.
2. Die zweite Interpretation basiert auf der Annahme, dass nicht beobachtete bzw. nicht beobachtbare Einflussgrössen die Effizienzunterschiede ausgleichen. Es ist z.B. denkbar, dass die Qualität der erbrachten Leistung in einzelnen Gruppen systematisch höher ist als bei den übrigen Beobachtungen, oder dass eine Untergruppe systematisch kompliziertere Fälle behandelt, d.h. solche, die mehr Ressourcen absorbieren. In beiden Fällen findet ein Selektionseffekt statt, der nicht berücksichtigt wird und entweder die Effizienzunterschiede verstärkt oder ausgleicht. Beiden Fällen ist gemeinsam, dass die Resultate der DEA nicht valide sind und eine zusätzliche Wahrscheinlichkeit eines Fehlers bei den Hypothesentests besteht.
3. Die dritte Interpretation ist, dass Gruppen vollständig oder teilweise unabhängige Abschnitte der effizienten Grenze definieren, wie in Kapitel 3 erläutert wurde. Sind die Abschnitte der GPM der einzelnen Gruppen voneinander unabhängig, können die Effizienzunterschiede zwischen den Gruppen nicht erkannt werden und die geschätzten Effizienzdifferenzen sind nicht bzw. bestenfalls zufällig unverzerrt.

¹¹Diese möglichen Ursachen beziehen sich nicht nur auf die Auswertung der Effizienz der Teilstichprobe der ineffizienten Beobachtungen. Auch die Zugehörigkeit zur Gruppe der effizienten Beobachtungen konnte im Rahmen des *Two-Part*-Modells nicht erklärt werden.

Die Folge ist, dass die Effizienzdifferenzen unterschätzt werden und die Wahrscheinlichkeit der Fehler zweiter Art zunimmt.

Welche dieser drei möglichen Interpretationen die „wahre“ ist, steht im Zentrum der folgenden Ausführungen. Es wird untersucht, ob Indizien für die zweite und dritte Interpretation existieren. Existieren solche, müssen die Folgerungen 5.1 revidiert werden. Existieren diese aber nicht, bietet sich die erste Interpretation als Ursache für den bescheidenen Erklärungsgehalt der ökonometrischen Auswertung an.

5.2 Bewertung der möglichen Ursachen der schwachen Erklärungskraft

In diesem Abschnitt soll die Frage geklärt werden, ob die bescheidene Erklärungskraft der ökonometrischen Auswertung der DEA-Effizienz der 310 Krankenhäuser (1) auf gruppenspezifisch unabhängige Abschnitte der GPM und verzerrte Effizienzdifferenzen oder (2) auf Unterschiede in der Qualität und dem *Case-Mix* der erbrachten Leistungen zurückzuführen ist.

5.2.1 Verzerrte Effizienzdifferenzen als mögliche Ursache

Wie in Abschnitt 3.4.1 erläutert wurde, bestehen dazu verschiedene Möglichkeiten.¹² Im folgenden Unterabschnitt werden die Referenzsets und die dazugehörigen ineffizienten Beobachtungen analysiert. Im darauf folgenden Unterabschnitt wird das in Abschnitt 3.4.2 vorgeschlagene Validitätsmass Δ berechnet und kommentiert. Schliesslich werden die Resultate beider Unterabschnitte miteinander verglichen.

Eine Analyse der Referenzsets

In einem ersten Schritt wird untersucht, in welchen Kombinationen effiziente Beobachtungen die GPM definieren.¹³ Dabei zeigt sich, dass die 37 effizienten Beobachtungen zusammen 145 Referenzsets für die 273 ineffizienten Beobachtungen bilden. Dann wird

¹²Auf eine Analyse der Korrelationen wird hier verzichtet.

¹³Die Informationen zu den Referenzsets als Grundlage dieser Analyse sind Bestandteil der Lösung der Linearen Programme.

überprüft, wie die Beobachtungen in Bezug auf ihre Gruppenzugehörigkeit auf die Referenzsets und die dazugehörenden ineffizienten Beobachtungen verteilt sind. Die hier interessierenden Gruppenmerkmale sind die Variablen, die die Anreizstrukturen im schweizerischen Krankenhaussektor repräsentieren sollen und in der ökonometrischen Analyse als Regressoren dienen.

Gruppenmerkmal Rechtsform Wird das Gruppenmerkmal „Rechtsform“ betrachtet, so zeigt sich, dass nur gerade 5 von 145 Flächen ausschliesslich durch öffentlich-rechtlich organisierte Beobachtungen definiert werden (2. und 4. Zeile in Tabelle 5.3). Zudem bilden diese nur für 11 ineffiziente Beobachtungen die Referenzsets und können daher als relativ unbedeutend bezeichnet werden. Die Beobachtungen mit privatrechtlicher Rechtsform definieren dagegen 74 Referenzflächen für insgesamt 151 ineffiziente Beobachtungen. Wie aus den beiden ersten Zeilen zu erkennen ist, treten 37 Referenzflächen mit insgesamt 54 ineffizienten Beobachtungen (d.h. 19.8%) in nicht-idealen Konstellationen auf: Sowohl alle Beobachtungen der Referenzsets wie auch alle ineffizienten Beobachtungen, die auf diese projiziert werden, weisen dasselbe Gruppenmerkmal auf und können somit keinen Beitrag zur Erklärung der gruppenspezifischen Effizienzunterschiede leisten. Dagegen tritt die ideale Konstellation in 42 Referenzsets für 108 ineffiziente Beobachtungen auf (39.6%): Die gruppenspezifische Ineffizienz wird vollumfänglich erkannt. Die übrigen 111 (40.7%) ineffizienten Beobachtungen werden auf Referenzsets projiziert, die sowohl von privatrechtlich wie auch öffentlich-rechtlich organisierten Beobachtungen definiert werden. Dies entspricht der in Abschnitt 3.3.3 erläuterten typischen Konstellation: Gruppenspezifische Ineffizienz kann teilweise erkannt werden. Eine Veränderung

Tabelle 5.3: Gruppenspezifische Abschnitte der GPM: Rechtsträger

| Konstellation | Referenz | Projektion | Anzahl Flächen | Anzahl Beobachtungen |
|----------------|----------|------------|----------------|----------------------|
| nicht-ideal | PRIV | PRIV | 33 | 49 |
| nicht-ideal | OEFF | OEFF | 4 | 5 |
| stabil-ideal | PRIV | | 41 | 102 |
| instabil-ideal | OEFF | | 1 | 6 |
| typisch | | | 66 | 111 |
| Total | | | 145 | 273 |

der gruppenspezifischen Ineffizienz der öffentlich-rechtlich organisierten Beobachtungen

müsste sich folglich in 39.6% der ineffizienten Beobachtungen vollständig und in 40.7% der Beobachtungen mindestens teilweise auf die durch die DEA ausgewiesene Effizienzdifferenz auswirken: Die durchschnittliche Effizienzdifferenz gemäss DEA müsste unterproportional zur Reduktion der gruppenspezifischen Ineffizienz zunehmen.

Gruppenmerkmal Subventionierung Als nächstes Gruppenmerkmal werden die subventionierten bzw. nicht subventionierten Beobachtungen untersucht. Aus Tabelle 5.4 ist ersichtlich, dass keine Referenzsets ausschliesslich durch subventionierte Beobachtungen definiert werden (vgl. 2. und 4. Zeile der entsprechenden Tabelle). Nicht subventionierte Beobachtungen stellen dagegen in 58 Fällen ein Referenzset für insgesamt 112 ineffiziente Beobachtungen dar. Über ein Drittel aller Referenzflächen (52, vgl. erste Zeile), auf die 98 ineffiziente Beobachtungen projiziert werden, stellen nicht-ideale Konstellationen dar. Nur gerade 6 Referenzflächen mit 14 ineffizienten Beobachtungen sind ideal. Den bedeutendsten Anteil bildet die typische Konstellation mit 87 Referenzflächen und 161 ineffizienten Beobachtungen. Bei 35.9% aller ineffizienten Beobachtungen wirkt sich damit

Tabelle 5.4: Gruppenspezifische Abschnitte: Subventionierung

| Konstellation | Referenz | Projektion | Anzahl Flächen | Anzahl Beobachtungen |
|----------------|----------|------------|----------------|----------------------|
| nicht-ideal | ohne | ohne | 52 | 98 |
| nicht-ideal | mit | mit | 0 | 0 |
| stabil-ideal | ohne | | 6 | 14 |
| instabil-ideal | mit | | 0 | 0 |
| typisch | | | 87 | 161 |
| Total | | | 145 | 273 |

eine Veränderung der gruppenspezifischen Ineffizienz nicht, bei 59.0% aller ineffizienten Beobachtungen nur teilweise auf die ausgewiesene Ineffizienz aus.

Gruppenmerkmal Defizitdeckung Aus Tabelle 5.5 ist ersichtlich, dass 86 der 145 Referenzsets ausschliesslich von Beobachtungen, die keine Defizitdeckung erhalten, definiert sind (vgl. 1. und 3. Zeile). Nur gerade 5 Referenzsets werden dagegen durch Beobachtungen mit Defizitdeckung definiert. Die nicht-ideale Konstellation tritt in 36 Referenzsets mit 48 ineffizienten Beobachtungen auf (17.6% aller ineffizienten Beobachtungen). Die ideale Konstellation wird bei 55 Referenzsets mit insgesamt 136 ineffizienten Beobachtungen

identifiziert. Die typische Konstellation umfasst 54 Referenzsets für insgesamt 89 ineffiziente Beobachtungen. Die Gefahr der unerkannten gruppenspezifischen Ineffizienz ist

Tabelle 5.5: Gruppenspezifische Abschnitte der GPM: Defizitdeckung

| Konstellation | Referenz | Projektion | Anzahl Flächen | Anzahl Beobachtungen |
|----------------|----------|------------|----------------|----------------------|
| nicht-ideal | ohne | ohne | 32 | 43 |
| nicht-ideal | mit | mit | 4 | 5 |
| stabil-ideal | ohne | | 54 | 130 |
| instabil-ideal | mit | | 1 | 6 |
| typisch | | | 54 | 89 |
| Total | | | 145 | 273 |

beim Gruppenmerkmal Defizitdeckung relativ klein. Immerhin in 49.8% aller ineffizienten Beobachtungen wird eine Veränderung der gruppenspezifischen Ineffizienz vollständig, und bei weiteren 89 Beobachtungen (32.6%) teilweise auf die ausgewiesene Ineffizienz übertragen. Nur bei 48 Beobachtungen (17.6%) ist die geschätzte Effizienz nicht von der gruppenspezifischen Ineffizienz abhängig.

Gruppenmerkmal Privatspital Tabelle 5.6 zeigt, dass 62 Referenzsets mit insgesamt 97 ineffizienten Beobachtungen (35.5%) in nicht-idealen Konstellationen auftreten. Dagegen werden nur 14 ineffiziente Beobachtungen auf 7 Referenzsets projiziert, die als ideal bezeichnet werden können. Der grösste Teil aller ineffizienten Beobachtungen (162, bzw. 59.3%) tritt in typischen Konstellationen auf und wird auf 76 verschiedene Referenzsets projiziert.

Tabelle 5.6: Gruppenspezifische Abschnitte der GPM: Privatspital

| Konstellation | Referenz | Projektion | Anzahl Flächen | Anzahl Beobachtungen |
|----------------|----------|------------|----------------|----------------------|
| nicht-ideal | nein | nein | 60 | 94 |
| nicht-ideal | ja | ja | 2 | 3 |
| stabil-ideal | nein | | 5 | 12 |
| instabil-ideal | ja | | 2 | 2 |
| typisch | | | 76 | 162 |
| Total | | | 145 | 273 |

Gruppenmerkmal Pauschalsubventionierung Aus Tabelle 5.7 ist ersichtlich, dass im Falle des Gruppenmerkmals Pauschalsubventionen (d.h. Subventionierung ohne komplementäre Defizitdeckung, vgl. S. 103) 55 Referenzsets mit insgesamt 103 ineffizienten Beobachtungen in nicht-idealen Konstellationen auftreten. Dagegen werden nur gerade 9 ineffiziente Beobachtungen auf 3 Referenzsets projiziert, die als ideal bezeichnet werden können. Der grösste Teil aller ineffizienten Beobachtungen (161 bzw. 59.0%) treten in typischen Konstellationen auf und werden auf 87 verschiedene Referenzsets projiziert. Bemerkenswert ist, dass keine Beobachtungen mit Pauschalsubventionen bei der Definition der GPM beteiligt sind. Dies ist erstaunlich, da die Vermutung geäußert wurde, dass diese Beobachtungen effizienter als die übrigen Beobachtungen sind.¹⁴

Tabelle 5.7: Gruppenspezifische Abschnitte der GPM: Pauschalsubvention

| Konstellation | Referenz | Projektion | Anzahl Flächen | Anzahl Beobachtungen |
|----------------|----------|------------|----------------|----------------------|
| nicht-ideal | mit | mit | 0 | 0 |
| nicht-ideal | ohne | ohne | 55 | 103 |
| stabil-ideal | mit | | 0 | 7 |
| instabil-ideal | ohne | | 3 | 2 |
| typisch | | | 87 | 161 |
| Total | | | 145 | 273 |

Folgerung 5.2 *Die Möglichkeit, Effizienzunterschiede entdecken zu können, scheint in Bezug auf das Gruppenmerkmal Defizitdeckung am günstigsten zu sein. Die idealen Konstellationen sind stark, die nicht-idealen dagegen schwach vertreten. Ähnlich günstig präsentiert sich die Situation für das Gruppenmerkmal Rechtsträger. Vergleichsweise schlecht schneidet die GPM in Bezug auf die Gruppenmerkmale Privatspital, Pauschalsubventionierung und Subventionierung ab. Der Anteil der nicht-idealen Konstellationen ist jeweils relativ gross, während die ideale Konstellation kaum auftritt.*

Folgerung 5.3 *Die Analyse der Referenzsets und der dazugehörenden ineffizienten Beobachtungen legt die Vermutung nahe, dass die geschätzten Effizienzdifferenzen in Bezug auf die fünf Gruppenmerkmale zu einem guten Teil unverzerrt sind.*

¹⁴Zu beachten ist, dass die Gruppe der pauschalsubventionierten Beobachtungen nur gerade 34 Beobachtungen aufweist und somit eher schwach in der Stichprobe vertreten ist.

Die Validität der Effizienzdifferenzen

Als Ergänzung zu den im vorigen Abschnitt gemachten qualitativen Aussagen bezüglich der Möglichkeiten, gruppenspezifische Ineffizienz entdecken zu können, soll nun das in Abschnitt 3.4.2 vorgeschlagene Validitätsmass als quantitativer Indikator dienen.¹⁵ Das Validitätsmass Δ wurde für die fünf Anreizvariablen gemäss Gleichung 3.23 berechnet, die Resultate sind in Tabelle 5.8 aufgeführt. In der ersten Spalte (*VRS*) finden sich die

Tabelle 5.8: Validitätsmasse Δ der Effizienzdifferenzen in Bezug auf die Anreizvariablen

| | <i>VRS</i> | <i>CRS^{CR}</i> |
|----------------------|------------|-------------------------|
| Rechtsträger | 0.342 | 0.773 |
| Subventionierung | 0.366 | 0.707 |
| Defizitdeckung | 0.386 | 0.804 |
| Privatspital | 0.246 | 0.560 |
| Pauschalsubventionen | 0.337 | 0.666 |

Validitätsmasse für die GPM mit variablen, in der zweiten Spalte (*CRS^{CR}*) diejenigen der GPM mit konstanten Skalenerträgen und restringierten Grenzkosten der Substitution und Transformation. Wie in Abschnitt 4.6 postuliert, nimmt die Möglichkeit, gruppenspezifisch unabhängige Abschnitte zu definieren, mit zunehmendem Grad der Restriktion der GPM ab: Für jedes Gruppenmerkmal nimmt das Validitätsmass Δ bei der DEA mit konstanten Skalenerträgen und restringierten Grenzkosten der Substitution und Transformation im Vergleich zur DEA mit variablen Skalenerträgen einen höheren Wert an.

Zur Bewertung der Schätzergebnisse sind die Validitätsmasse der zweiten Spalte relevant, da die relative Distanz zur *CRS^{CR}*-Effizienzgrenze als abhängige Variable der ökonomischen Auswertung gewählt wurde. Es zeigt sich, dass das Validitätsmass Werte zwischen 0.56 und 0.804 annimmt. Am günstigsten schneidet dabei das Gruppenmerkmal Defizitdeckung mit $\Delta = 0.804$ ab. Ebenfalls eine hohe Validität der Effizienzdifferenzen kann dem Indikator des Rechtsträgers ($\Delta = 0.773$) zugeschrieben werden. Eine vergleichsweise geringe Validität der Effizienzdifferenzen ist dem Gruppenmerkmal Privatspital sowie Pauschalsubventionen mit einem Δ von 0.56 bzw. 0.67 zuzuschreiben.

¹⁵Im dritten und vierten Kapitel wurde gezeigt, dass Werte von Δ um eins mit validen Effizienzdifferenzen in Verbindung zu bringen sind. Mit abnehmenden Werten des Validitätsmasses steigt die Gefahr verzerrter Effizienzdifferenzen und auch der erwartete Anteil der Fehler zweiter Art bei den Hypothesentests.

Diese Ergebnisse stimmen gut mit den Folgerungen 5.2 und 5.3 überein. Auch dort erschienen die Defizitdeckung und der Rechtsträger in besonders günstigem Licht, während die übrigen Gruppenmerkmale vergleichsweise schlecht abschnitten. Die Erkenntnisse des Kapitels 4 zeigen aber auch, dass bereits in diesem Bereich der Δ s mit einer erhöhten Wahrscheinlichkeit von Fehlern zweiter Art zu rechnen ist (vgl. Tabelle 4.9). Die Erkenntnisse des vierten Kapitels basieren allerdings auf einer Transformationsfunktion mit zwei Inputs und einem Output, womit der zu erwartende Anteil der Fehler zweiter Art nicht auf die hier präsentierte Transformationsfunktion mit jeweils fünf Inputs und Outputs übertragen werden darf. Trotzdem liegt der Schluss nahe, dass die Effizienzunterschiede in Bezug auf die Anreizvariablen in der DEA tendenziell unterschätzt werden und die Gefahr von Fehlern zweiter Art leicht erhöht ist.

Diese Erkenntnisse lassen den Schluss zu, dass die von der DEA ausgewiesenen Effizienzdifferenzen recht valide sind und die mangelnde Signifikanz nicht das Ergebnis von gruppenspezifisch unabhängigen Abschnitten der GPM ist (vgl. Tabelle 5.2).

Zusätzliche Informationen zu den Eigenschaften der GPM liefert die sogenannte *Free-Disposal-Hull* (FDH).¹⁶ Bei diesem Verfahren wird der Input-Outputvektor jeder Beobachtung mit allen übrigen Input-Outputvektoren verglichen. Es wird untersucht, ob der zu evaluierende Vektor andere dominiert oder selbst dominiert wird (vgl. zur Illustration Abbildung 2.6 auf S. 15). Es hat sich gezeigt, dass von den 310 Beobachtungen nur gerade deren 12 von 17 Beobachtungen dominiert werden. Die überwältigende Mehrheit der Stichprobe (281 Beobachtungen) werden weder dominiert, noch dominieren sie selbst: Gemäss FDH sind diese 281 Beobachtungen effizient. Dieses Resultat ist ein weiteres Indiz dafür, dass die Beobachtungen eher „nebeneinander“ auf den Isoquanten als „hintereinander“ liegen.

Folgerung 5.4 *Die GPM der hier analysierten Stichprobe von 310 Beobachtungen des schweizerischen Krankenhaussektors besteht kaum aus gruppenspezifisch unabhängigen Abschnitten. Dies zeigt eine Auswertung der Referenzsets und der dazugehörigen ineffizienten Beobachtungen und die Werte der Validitätsmasse Δ . Zudem ist die GPM durch wenige dominierende und noch weniger dominierte Beobachtungen gekennzeichnet. Diese Erkenntnisse lassen den Schluss zu, dass die Effizienzunterschiede gemäss DEA valide sind.*

¹⁶An dieser Stelle wird nicht detailliert auf dieses Verfahren, das von Deprins, Simar und Tulkens (1984) entwickelt wurde, eingegangen. Die FDH ist ein nichtparametrisches Distanzmass, bei der die drei Konvexitätsannahmen der Technologiemenge fallen gelassen werden (vgl. Abschnitt 2.1.2).

5.2.2 Unterschiede in der Qualität und im Case-Mix als mögliche Ursache

Als zweite mögliche Ursache für den schlechten Erklärungsgehalt der ökonometrischen Auswertung der DEA-Effizienz kommen die nicht erfasste Qualität und auch Unterschiede im *Case-Mix* in Frage, die – je nach Umständen – die Effizienzunterschiede verstärken oder abschwächen. Leider liegen für diese Stichprobe weder Informationen zur Qualität der erbrachten Leistungen noch diagnostische Informationen zur Beurteilung der Komplexität der behandelten Fälle vor. Allerdings wurde in der ökonometrischen Auswertung der DEA-Effizienz mit den Kontrollvariablen PRIV% (für die Qualität) und IMEXQ (unter anderem für den *Case-Mix*) der Versuch unternommen, für solche Unterschiede zu kontrollieren. Ein plausibles Argument für geringe Unterschiede in Bezug auf die Qualität der erbrachten Leistungen – und somit gegen die zweite in Abschnitt 5.1.6 aufgeführte Ursache – dürfte zusätzlich in der direktdemokratischen Struktur der Schweiz zu finden sein. Es ist kaum denkbar, dass wahrnehmbar schlechter erbrachte Leistungen der Krankenhäuser einer Ortschaft bzw. einer Region von der Bevölkerung bzw. von den Wählern hingenommen werden. Grundsätzlich hat die Bevölkerung zwei Optionen: Sie kann sich in einer anderen Region (mit qualitativ hochwertigeren Krankenhäusern) behandeln lassen (*Exit*) oder durch ihr Wahlverhalten dafür sorgen, dass die Qualität verbessert wird (*Voice*). Die erste Option hätte zur Folge, dass die Patienten sich in der Region behandeln lassen, in der die Qualität ihren Präferenzen am besten entspricht: Es kommt zu regional unterschiedlichen Qualitätsniveaus. Da die Mobilität der Patienten eingeschränkt ist – diese wird (interkantonal) nur durch Zusatzversicherungen oder einen ärztlichen Entscheid möglich – bleibt den Patienten die Option *Voice*, die dazu führt, dass die Qualitätsunterschiede tendenziell abgebaut werden.

Folgerung 5.5 *Ob Qualitätsunterschiede und der Case-Mix für die mangelnde Erklärungskraft verantwortlich sind, kann im Rahmen dieser Arbeit nicht schlüssig geklärt werden. Zu beachten ist, dass die Effizienzdifferenzen nur unter bestimmten Annahmen durch Unterschiede in der Qualität bzw. dem Case-Mix vermindert werden und dass bei der ökonometrischen Auswertung der relativen Distanz zur GPM – mindestens approximativ – für diese Grössen kontrolliert wurde.*

5.3 Fazit zur DEA im schweizerischen Krankenhaussektor

Dieses Kapitel hat gezeigt, dass die Effizienzunterschiede im schweizerischen Krankenhaussektor nur teilweise auf die Anreizstrukturen zurückzuführen sind. Die geschätzten Koeffizienten sind im Absolutwert klein und bestenfalls knapp signifikant von null verschieden. Es konnte zudem gezeigt werden, dass einiges dafür spricht, dass die Effizienzdifferenzen valide sind. Somit kann die dritte in Abschnitt 5.1.6 genannte mögliche Ursache für den mangelnden Erklärungsgehalt der ökonometrischen Analyse der Effizienz ohne grosses Risiko ausgeschlossen werden. Die zweite Interpretation kann im Rahmen dieser Arbeit nicht überprüft werden. Allerdings sprechen – wie in Abschnitt 5.2.2 erläutert wurde – gute Argumente dafür, dass nicht Qualitätsunterschiede bzw. verschiedene *Case-Mix* der erbrachten Leistungen für die erheblichen Effizienzdifferenzen verantwortlich sind.

Somit bleibt die Erklärung Nr. 1 übrig: Die Anreizstrukturen im schweizerischen Krankenhaussektor sind derart schwach, dass sich die Entscheidungsträger nicht gezwungen sehen, die Leistung effizient zu erbringen. Zwei mögliche Ursachen für die mangelnden Anreizstrukturen sind denkbar:

1. Mechanismen zur Bewertung der Effizienz, die eine Reduktion der asymmetrischen Informationen ermöglichen, sind eine Grundvoraussetzung für wirkungsvolle Anreizstrukturen. Versagen diese Mechanismen, ist die Leistungsbewertung und die Sanktionierung ineffizienter Leistungserbringer nicht möglich.
2. Die Regulierungsbehörden verzichten auf die Sanktionierung der ineffizienten Leistungserbringer, obwohl die Bewertung der Leistungserbringung grundsätzlich möglich wäre.

Zu Punkt eins ist zu sagen, dass in der Schweiz im politischen Prozess weder ein Konsens in Bezug auf die Definition des Leistungskataloges der Krankenhäuser noch die zu verwendende Methode zur Bewertung der Effizienz besteht. Somit ist die Sanktionierung ineffizienter Leistungserbringer im schweizerischen Krankenhaussektor *a priori* zum scheitern verurteilt. Auch wenn die Grundvoraussetzungen für die Sanktionierung (momentan noch) nicht erfüllt sind,¹⁷ könnte die Ursache für die schwachen Anreizstrukturen auch

¹⁷Vielleicht wird sich der Druck auf die Politiker infolge der sog. Kostenexplosion in Zukunft noch verstärken, so dass die Leistungsbewertung gesetzlich vorgeschrieben wird. In Deutschland z.B. ist der

im zweiten Punkt zu suchen sein. Der Verzicht der Sanktionierung könnte auf einem gesellschaftlichen Konsens beruhen: Die Mehrheit der Wähler wünscht, dass Krankenhäuser nicht sanktioniert werden. In diesem Fall ist die Effizienz offensichtlich sekundär. Auch wenn dies nicht zutrifft, die Mehrheit der Wähler also eine Sanktionierung zur Steigerung der Effizienz wünscht, ist es denkbar, dass die Politiker die Wählerpräferenzen infolge asymmetrischer Informationen oder nicht bindender Wiederwahlrestriktion nicht berücksichtigen. Eine weitere Interpretation zum zweiten Punkt lautet dahingehend, dass die Regulierungsbehörde selbst keine Anreize hat, die ineffizienten Leistungserbringer zu sanktionieren. Gerade wenn die Sanktionierung mit Kosten verbunden ist – eine plausible Annahme unter Berücksichtigung des zu erwartenden Widerstandes seitens der Entscheidungsträger und Arbeitnehmer der zu sanktionierenden Krankenhäuser – und die Regulierungsbehörde infolge asymmetrischer Information keine Konsequenzen im Falle unterlassener Sanktionierung zu gewärtigen hat, ist nicht mit einer rigorosen Leistungsbewertung und Sanktionierung zu rechnen. In dieser Situation liegt ein verschachteltes Prinzipal-Agenten-Problem vor, bei dem nicht nur ein Informationsvorsprung zwischen Leistungserbringer und Regulierungsbehörde besteht, sondern zudem die Regulierungsbehörde als Agent der Wähler ihrerseits über einen Informationsvorsprung verfügt. Somit stellt sich die Frage, ob die Regulierungsbehörde ein geeigneter Sachwalter sein kann oder ob sie eher die Interessen der Leistungserbringer vertritt [vgl. Laffont und Tirole (1993), Kap. 11]. Über die Ursachen der schwachen Wirkung der Anreizstrukturen kann hier nur spekuliert werden. Es besteht aber unabhängig von den wahren Ursachen die erhöhte Gefahr, dass die effizienten Leistungserbringer die ineffizienten imitieren, die Produktivität insgesamt sinkt und knappe Ressourcen verschwendet werden.

Folgerung 5.6 *Die relativ hohen Werte der Validitätsmasse der Effizienzdifferenzen in Bezug auf die Anreizvariablen legen den Schluss nahe, dass im schweizerischen Krankenhaussektor die Effizienzunterschiede kaum verzerrt und folglich nicht auf die Anreizstrukturen zurückzuführen sind.*

Zum Abschluss dieses Kapitels bleibt anzumerken, dass die hier verwendete Stichprobe in vielerlei Hinsicht Wünsche offen lässt. Ideal wäre vor allem, wenn Beobachtungen über einen längeren Zeitraum vorlägen, um die zeitliche Entwicklung besser verfolgen

Krankenhausbetriebsvergleich seit In-Kraft-Treten der Bundespflegesatzverordnung (BPflV) 1995 gesetzlich verankert.

zu können. Dies würde vermutlich auch die Chancen erhöhen, dass weniger Kompromisse bei den ökonometrischen Schätzmethoden gemacht werden müssen. Ausserdem wäre interessant, den Vergleich mit ausländischen Krankenhäusern anzustellen, der einen Rückschluss auf die Auswirkungen der unterschiedlichen Gesundheitssysteme auf die Effizienz ermöglichte. In Bezug auf die Qualität der Daten ist zu erwähnen, dass zusätzliche krankenhauspezifische diagnostische Informationen die Aussagekraft der Analyse wesentlich erhöhen würde.

Kapitel 6

Schlussfolgerungen

In dieser Arbeit wurden die statistischen Eigenschaften der *Data Envelopment Analysis* (DEA) untersucht. Im Vordergrund stand die Frage, ob und unter welchen Umständen die DEA in empirischen Anwendungen konsistente Schätzer der Effizienz liefert und damit valide Hypothesentests ermöglicht. Die statistischen Eigenschaften sind gerade bei volkswirtschaftlichen Anwendungen – bei denen oft regulierte Märkte analysiert werden – von herausragender Bedeutung, da die Konsequenzen von fehlerhaften Entscheiden gravierend sein können und eine nachträgliche Anpassung, gerade in demokratischen Gesellschaften, mit nicht zu unterschätzenden Transaktionskosten verbunden sein dürften.

Nachdem die ökonomischen Effizienzkonzepte erläutert wurden, erfolgte eine Einführung in die nichtparametrischen Methoden der Effizienzmessung. Gemein ist diesen Methoden, dass eine Effizienzgrenze ermittelt wird, an der die ineffizienten Beobachtungen gemessen werden. Die Distanz einer ineffizienten Beobachtung zur effizienten Grenze wird dann als *technische Effizienz* interpretiert. Die nichtparametrischen Methoden zur Effizienzmessung erlauben aber auch die Ermittlung der *Skaleneffizienz* und der *allokativen Effizienz*. Während die *Skaleneffizienz* durch eine Anpassung der Strukturannahmen der Technologie ermittelt werden kann, bedarf es zur Bestimmung der *allokativen Effizienz* zusätzlicher Informationen über die Preise der Produktionsfaktoren und der produzierten Güter. Ausserdem wurde untersucht, wie unterschiedliche Rahmenbedingungen in die Effizienzmessung integriert werden können. Es ist offensichtlich, dass die Rahmenbedingungen den Produktionsprozess beeinflussen und es besteht ein Konsens, dass diese nicht unberücksichtigt bleiben dürfen. Grundsätzlich existieren zwei Alternativen: Einerseits können die Rahmenbedingungen direkt in die Effizienzmessung integriert werden. Diese

Vorgehensweise wird als *One-Stage*-Ansatz bezeichnet. Andererseits kann die Korrektur nach erfolgter Effizienzmessung erfolgen. Bei diesem sogenannten *Two-Stage*-Ansatz soll die Distanz zur effizienten Grenze (die Resultate der Effizienzmessung) in der zweiten Stufe in einem ökonometrischen Modell durch die Rahmenbedingung erklärt werden. Die Distanz zur effizienten Grenze wird bei dieser Variante in zwei Komponenten zerlegt: In die Wirkung der Rahmenbedingungen und in reine Ineffizienz. Der *Two-Stage*-Ansatz hat sich im Zusammenhang mit der DEA in der Praxis durchgesetzt.

Anschliessend wurden die statistischen Eigenschaften der Schätzer der DEA-Effizienz diskutiert. Banker (1993) hat gezeigt, dass die DEA verzerrte, aber unter bestimmten Bedingungen konsistente Schätzer der Effizienz liefert. Eine explizite Bedingung bezieht sich auf die Effizienzverteilung, eine implizite auf die Verteilung der Beobachtungen in der Technologiemenge. Beide Bedingungen erweisen sich als höchst problematisch, da sie sich in einer empirischen Anwendung nicht überprüfen lassen. Der Forscher wird auch nach erfolgter DEA nicht wissen, wie stark die Schätzer verzerrt bzw. ob sie überhaupt konsistent sind. Rückschlüsse aufgrund solcher Schätzer auf die Allgemeinheit, wie dies bei Hypothesentests üblich ist, sind somit möglicherweise fehlerhaft. Zur Illustration dieser Problematik wurden verschiedene Konstellationen für eine Technologie mit zwei Produktionsfaktoren und einem Gut grafisch erläutert. Mit der Konstellation wird dabei die Verteilung der zu analysierenden Firmen, Unternehmungen usw. in der Technologiemenge beschrieben. Es hat sich gezeigt, dass in *nicht-idealen* Konstellationen, bei denen die Firmen bezüglich der Input- und Outputverhältnisse *systematisch* in der Technologiemenge verteilt sind, die Schätzer der Effizienz unterschiedlich stark verzerrt sein können und folglich auch die Effizienzdifferenzen sich nicht valide ermitteln lassen. In *idealen* Konstellationen, bei denen die Firmen bezüglich der Input- und Outputverhältnisse unabhängig verteilt sind, werden die Effizienzdifferenzen dagegen valide geschätzt. Diese Erkenntnisse konnten auch mathematisch für den Fall mit beliebig vielen Produktionsfaktoren und Gütern bestätigt werden. Problematisch ist, dass die Konsistenzbedingungen im Normalfall dem ökonomischen Handeln der Produzenten widersprechen. Unterschiedliche Rahmenbedingungen führen dazu, dass sich das Verhalten der Produzenten – d.h. die Wahl der Produktionsfaktoren und Güter – unterscheidet. Stehen die Rahmenbedingungen in einem systematischen Zusammenhang mit dem Verhalten, ist die implizite Konsistenzbedingung der DEA (die Beobachtungen sind unabhängig und identisch verteilt) verletzt und es besteht die Gefahr, dass die Effizienzdifferenzen verzerrt sind, d.h. nicht valide ermittelt werden können.

Diese Erkenntnisse führen zu einer Erweiterung der Konsistenzbedingungen der DEA: Die Schätzer der Effizienz sind unabhängig von der Verteilung der Firmen in der Technologiemenge konsistent, wenn die Wahrscheinlichkeit einer fast vollständig effizienten Beobachtung für jede Teilmenge der Grundgesamtheit, die eine Technologie (d.h. einen Abschnitt der GPM) definiert, grösser null ist.

Bei finiten Stichproben sind die Schätzer infolge der mangelnden Erwartungstreue verzerrt. Werden Effizienzunterschiede zwischen Gruppen von Beobachtungen analysiert, müssen die Verzerrungen für die zu analysierenden Gruppen von Beobachtungen im Erwartungswert gleich gross sein. Trifft dies zu, sind die geschätzten Effizienzdifferenzen unverzerrt. Ist dies dagegen nicht der Fall (oder kann die Effizienz nur für eine Teilstichprobe konsistent geschätzt werden) sind die geschätzten Effizienzdifferenzen nicht valide und die Wahrscheinlichkeit von Fehlern bei den Hypothesentests steigt. Eine Bedingung für unverzerrte Effizienzdifferenzen lautet daher: Die geschätzte Effizienz ist im Erwartungswert gleich stark verzerrt, wenn die Wahrscheinlichkeit einer fast vollständig effizienten Beobachtung für jede Teilmenge der Grundgesamtheit, die eine Technologie (einen Abschnitt der GPM) definiert, gleich gross ist.

Diese Erkenntnisse sind in der empirischen Forschung allerdings von beschränktem Wert, da sie sich nicht überprüfen lassen. Daher wurde eine Kennzahl entwickelt, die eine Beurteilung der Validität der geschätzten Effizienzdifferenzen zulässt. Das *Validitätsmass der Effizienzdifferenzen* setzt sich aus zwei Kennzahlen zusammen, die jeweils angeben, in welchen Ausmass eine Veränderung der Daten *einer* Teilstichprobe sich auf die durchschnittliche geschätzte Effizienz *beider* Teilstichproben auswirkt. Das Validitätsmass gibt Auskunft darüber, wie sich diese Änderung der Daten auf die geschätzte *Effizienzdifferenz* auswirkt. Im Idealfall nimmt das Validitätsmass einen Wert von eins an. Dies bedeutet, dass die Veränderung der Daten sich zu 100 Prozent auf die geschätzte Effizienzdifferenz auswirkt. Ein Wert von null zeigt andererseits an, dass die Veränderung der Daten keine Konsequenzen auf die Effizienzdifferenz hat. Das Validitätsmass wurde anhand eines dichotomen Gruppenmerkmals illustriert. Allerdings kann es für jedes diskrete Gruppenmerkmal mit beliebig vielen Teilstichproben berechnet werden. Die Validität der Effizienzdifferenzen bezüglich einer stetigen Variable lässt sich berechnen, indem die entsprechende Variable in Kategorien eingeteilt, d.h. in ein diskretes Messniveau überführt wird.

In einer *Monte-Carlo-Simulation* wurden folgende Hypothesen überprüft:

Hypothese 1 *Nimmt das Validitätsmass in Bezug auf ein Gruppenmerkmal einen Wert*

um eins an, sind die erwarteten Effizienzdifferenzen zwischen den Gruppen valide, d.h. unverzerrt.

Hypothese 2 *Bestehen Effizienzunterschiede zwischen zwei zu evaluierenden Gruppen, ist bei niedriger Validität der geschätzten Effizienzdifferenzen mit einer erhöhten Wahrscheinlichkeit von Fehlern zweiter Art (Nichterkennen tatsächlicher Effizienzunterschiede) zu rechnen.*

Untersucht wurde, unter welchen Umständen unterschiedlich starke Verzerrungen bei der Schätzung der Effizienz zwischen den Teilstichproben auftreten und wie sich diese auf die Ebene der Hypothesentests übertragen. Im Vordergrund stand die Frage, ob das Validitätsmass ein geeigneter Indikator für das Ausmass der Verzerrungen der Effizienzdifferenzen ist. In der *Monte-Carlo*-Simulation wurde eine *Cobb-Douglas*-Produktionsfunktion mit zwei Produktionsfaktoren und konstanten Skalenerträgen unterstellt. Die Auswertung ergab, dass Parameterkombinationen, die mit hohen Werten des Validitätsmasses einhergehen, zu unverzerrten und somit validen Effizienzdifferenzen führten. In allen Konstellationen mit kleineren Werten des Validitätsmasses ist die geschätzte Effizienz bei der ineffizienteren Gruppe nach oben verzerrt. Das Ausmass der Verzerrung nimmt im Erwartungswert mit abnehmenden Werten des Validitätsmasses *ceteris paribus* zu. Da die geschätzte Effizienz bei der effizienteren Gruppe dagegen beinahe unverzerrt ist, sind die Effizienzdifferenzen nicht valide. Des Weiteren hat sich gezeigt, dass die Hypothesentests gute Resultate liefern: Der Anteil der Fehler erster Art entspricht – dem gewählten Signifikanzniveau entsprechend – rund 5 Prozent. Der Anteil der Fehler zweiter Art nimmt mit zunehmenden Werten des Validitätsmasses ab und entspricht ebenfalls den theoretischen Erwartungen. Diese Resultate stützen die beiden formulierten Hypothesen und führen zur Folgerung, dass ein hoher Wert des Validitätsmasses eine hinreichende Bedingung für valide Effizienzdifferenzen ist.

Zur Illustration des Validitätsmasses wurde eine Anwendung der DEA im schweizerischen Krankenhaussektor präsentiert. Die ökonometrische Auswertung der DEA-Effizienz hat dabei ergeben, dass die Effizienzdifferenzen kaum auf die Anreizstrukturen¹ zurückzuführen sind. Drei mögliche Gründe für die schwache Erklärungskraft wurden aufgeführt. Erstens besteht die Möglichkeit, dass die Effizienzdifferenzen tatsächlich nicht auf die Anreizstrukturen zurückzuführen sind. In dieser Situation ist die Interpretation der geschätz-

¹Die Anreizstrukturen wurden durch die Dummyvariablen Subventionierung, Defizitdeckung, Rechtsträger, Privatspital und Pauschalsubventionierung erfasst.

ten Koeffizienten korrekt. Die Frage, wieso die Anreizstrukturen nicht wirken, bleibt aber unbeantwortet. Eine mögliche Ursache könnte sein, dass effiziente Krankenhäuser die ineffizienten imitieren und nur bereit sind, einen Effort zur Steigerung der Effizienz zu leisten, wenn sie dafür entschädigt werden. Eine zweite Ursache der schwachen Erklärungskraft könnte sein, dass Qualität und Komplexität der erbrachten Leistungen positiv mit der wahren (unbekannten) Effizienz korreliert sind und sowohl Qualität wie auch Komplexität *ceteris paribus* mehr Ressourcen absorbieren. Treffen beide Annahmen zu, würden die Krankenhäuser mit qualitativ hochwertigen und komplexen Leistungen zu unrecht als relativ ineffizient eingeschätzt werden, da die DEA die Qualität und die Komplexität innerhalb der definierten Outputkategorien nicht berücksichtigt. Die Effizienzdifferenzen würden folglich *ceteris paribus* unterschätzt und bei der Interpretation der geschätzten Koeffizienten stiege die Wahrscheinlichkeit eines Fehlers zweiter Art. Als dritte Ursache wurde die mangelnde Validität der geschätzten Effizienzdifferenzen in Bezug auf die Anreizvariablen genannt, was verzerrte, d.h. unterschätzte Effizienzdifferenzen nach sich ziehen würde. Falls keine Indizien für die zweite oder dritte Begründung existieren, ist die erste am wahrscheinlichsten.

In Bezug auf die zweite Erklärung wurde argumentiert, dass Indikatoren der Komplexität der behandelten Fälle und der Qualität in die Schätzgleichung integriert wurden. Zudem spricht die Möglichkeit der direktdemokratischen Einflussnahme dagegen, dass substantielle interregionale Qualitätsunterschiede entstehen können. Die Analyse der Referenzsets und auch die berechneten Validitätsmasse lassen vermuten, dass die Daten für eine statistische Auswertung geeignet sind. Die Validitätsmasse nehmen Werte zwischen rund 0.5 und 0.8 an, wobei die beiden Variablen Defizitdeckung und Rechtsträger am besten abschneiden. Bei den Variablen Privatspital und Pauschalsubventionierung, mit einem Validitätsmass von 0.56 bzw. 0.67, besteht allerdings die Gefahr, dass die Effizienzdifferenzen unterschätzt wurden. Die dritte Erklärung für die schwache Erklärungskraft kann somit nicht definitiv ausgeschlossen werden, obwohl einiges dafür spricht, dass die Effizienzdifferenzen valide sind. Somit besteht die berechtigte Vermutung, dass die Effizienzunterschiede tatsächlich nicht auf die Anreizstrukturen zurückzuführen sind.

Im Bereich der statistischen Fundierung der nichtparametrischen Methoden der Effizienzmessung wurden in den vergangenen Jahren enorme Fortschritte erzielt. Die von Banker (1993) explizit formulierten Bedingungen für die Konsistenz setzen allerdings eine implizite Bedingung voraus, die in der empirischen Forschung nicht *a priori* als erfüllt betrachtet werden darf. Diese Arbeit hatte zum Ziel, die Konsistenz der DEA im Kontext

der empirischen Forschung zu analysieren. Die Resultate basieren auf der Simulation einer Produktionsfunktion mit zwei Produktionsfaktoren und einem Gut. Die DEA wird in der Praxis aber häufig zur Analyse von komplexen Produktionsprozessen mit mehreren Inputs und Outputs verwendet. Eine zweite Restriktion stellte zudem die Tatsache dar, dass die Rahmenbedingungen ausschliesslich die Unterteilung der Beobachtungen in zwei Gruppen beinhaltete. In der Simulation konnten die geschätzten Effizienzdifferenzen daher mittels T-Tests auf ihre Signifikanz überprüft werden. In empirischen Anwendungen existieren aber i.d.R. mehrere Faktoren, die zu unterschiedlichen Rahmenbedingungen führen. Interessant wäre daher der Einschluss mehrerer diskreter und/oder stetiger Rahmenbedingungen in die Simulation, was eine Auswertung der simulierten DEA-Resultate mittels multipler Regressionsanalyse oder des *Tobit*-Schätzverfahrens nahe legen würde. Eine dritte Einschränkung bestand darin, dass die geschätzten Effizienzverteilungen ausschliesslich bezüglich des Mittelwertes ausgewertet wurden und die Streuung, der zweite Parameter zur Beschreibung der Normalverteilung, unberücksichtigt blieb.

Monte-Carlo-Simulationen ohne diese Einschränkungen stellen eine Herausforderungen an die zukünftige Forschung dar. Es ist nicht anzunehmen, dass solche Simulationen zu qualitativ verschiedenen Aussagen führen werden. Allerdings würde weitere *Monte-Carlo*-Evidenz exakte quantitative Aussagen zur Bewertung der Resultate bezüglich der Fehlerwahrscheinlichkeiten bei Hypothesentests empirischer Arbeiten ermöglichen.

Anhang A

MATLAB *Monte-Carlo*-Syntax (©Lukas Steinmann)

```
%%%%%%%% Start der Monte-Carlo-Simulation
randn('state',(sum(clock)+sum(date))^6);
varr=125;
vars=6;

varq=5;

varv=3;

varg=3;

%%%%%%%% Definieren der Vektoren
output=zeros(1,73); output1=zeros(1,73); output2=zeros(1,73);
output3=zeros(1,73); output4=zeros(1,73);

%%%%%%%% Start der Monte-Carlo-Simulation
%%%%%%%% Äussere Schleifen: zu verändernde Parameterwerte
timall0= cputime;                                % Startzeit speichern
for g=1:varg
    gg=g-1;
    for v=1:varv
        vv=v-1;
        for q=1:varq
            qq=q-1;
            for s=1:vars
                ss=s-1;

                %%%%%%%%% Innere Schleife: Iterationen
                for r=1:varr;
                    time0= cputime;                % Startzeit der Iteration r

                    %%%%%%%%% Festlegen der fixen Simulationsparameter
                    ga=100;
                    gb=100;
                    n=ga+gb;
                    ea=0.85;
                    eb=0.85;
                    sia=0.15+gg*0.05;
                    sib=0.15+gg*0.05;
                    psia=qq*0.05;
                    psib=0;
                    psiu=1/1.1;
                    psid=1/0.9;
                    a=0.5;

                    % Gruppengrösse A
                    % Gruppengrösse B
                    % Stichprobengrösse
                    % Erwartungswert Effizienz Gruppe A
                    % Erwartungswert Effizienz Gruppe B
                    % Standardabweichung der Effizienzverteilung der Gruppe A
                    % Standardabweichung der Effizienzverteilung der Gruppe B
                    % Gruppenspezifische Ineffizienz Gruppe A
                    % Gruppenspezifische Ineffizienz Gruppe B
                    % Parameter „Psi“ für Sensitivitätstest
                    % Parameter „Psi“ für Sensitivitätstest
                    % Exponent der Cobb-Douglas-Funktion Input 1
```

```

exa=50+6*ss;
exb=50-6*ss;

sxa= 10-vv*3;
sxb= 10-vv*3;
eya=100;
eyb=100;
sya=20;
syb=20;

%%%%%% Berechnen der Input-Outputdaten
%%%%%% Generieren der Inputmixe
xa=zeros(ga,1);
xa=normrnd(exa,sxa,ga,1)./100;
for i=1:ga;
    if xa(i)<=0 xa(i)=0.01;
    elseif xa(i)>=1 xa(i)=0.99;
    end;
end;
xb=zeros(gb,1);
xb=normrnd(exb,sxb,gb,1)./100;
for j=1:gb;
    if xb(j)<=0 xb(j)=0.01;
    elseif xb(j)>=1 xb(j)=0.99;
    end;
end;

%%%%%% Generieren des Effizienzvektors
ineffa=zeros(ga,1);
ineffa=normrnd(ea,sia,ga,1);
for k=1:ga;
    if ineffa(k)>1 ineffa(k)=1;
    elseif ineffa(k)<=psia ineffa(k)=ea;
    end;
end;
ineffa=ineffa-psia;

ineffb=zeros(gb,1);
ineffb=normrnd(eb,sib,gb,1);
for m=1:gb;
    if ineffb(m)>1 ineffb(m)=1;
    elseif ineffb(m)<=psib ineffb(m)=eb;
    end;
end;
ineffb=ineffb-psib;

%%%%%% Berechnen der Outputs
ya=zeros(ga,1);
ya=normrnd(eya, sya, ga,1);
for i=1:ga;
    if ya(i)<=0 ya(i)=1;
    end;
end;
yb=zeros(gb,1);
yb=normrnd(eyb, syb, gb,1);
for j=1:gb;
    if yb(j)<=0; yb(j)=1;
    end;
end;
xa2=zeros(ga,1);
xa1=zeros(ga,1);

% Erwartungswert Anteil Input 1
% zur Summe beider Inputs Gruppe A
% Erwartungswert Anteil Input 1
% zur Summe beider Inputs Gruppe B
% Standardabweichung Inputverhältnis Gruppe A
% Standardabweichung Inputverhältnis Gruppe B
% Erwartungswert Output Gruppe A
% Erwartungswert Output Gruppe B
% Standardabweichung Output Gruppe A
% Standardabweichung Output Gruppe B

% Vektor definieren bzw. reseten
% Generiert den Vektor der Inputmixe der Gruppe A

% Inputmixe  $\leq 0$  auf 0.01 setzten
% Inputmixe  $\geq 1$  auf 0.99 setzten

% Vektor definieren bzw. reseten
% Generiert den Vektor der Inputmixe der Gruppe B

% Inputmixe  $\leq 0$  auf 0.01 setzten
% Inputmixe  $\geq 1$  auf 0.99 setzten

% Vektor definieren bzw. reseten
% Generiert den Effizienzvektor der Gruppe A

% Keine Elemente des Effizienzvektors  $> 1$ 
% Keine Elemente m des Effizienzvektors  $< psia$ 

% Integration der gruppenspezifischen Ineffizienz

% Vektor definieren bzw. reseten
% Generiert den Effizienzvektor der Gruppe B

% Keine Elemente m des Effizienzvektors  $> 1$ 
% Keine Elemente m des Effizienzvektors  $< psib$ 

% Integration der gruppenspezifischen Ineffizienz (psib=0)

% Vektor definieren bzw. reseten
% Generiert den Outputvektor der Gruppe A

% Keine Elemente des Outputvektors sind  $\leq 0$ 

% Vektor definieren bzw. reseten
% Generiert den Outputvektor der Gruppe B

% Keine Elemente des Outputvektors sind  $\leq 0$ 

% Vektor definieren bzw. reseten
% Vektor definieren bzw. reseten

```

```

xa2=ya./(xa./(ones(ga,1)-xa)).^a;
% Berechnen des 2. Inputs der Gruppe A,
% gegeben Output und Inputmix
xa1=(xa./(ones(ga,1)-xa)).*xa2;
% Berechnen des 1. Inputs der Gruppe A,
% gegeben Output und Input 2

xb2=zeros(gb,1);
% Vektor definieren bzw. reseten
xb1=zeros(gb,1);
% Vektor definieren bzw. reseten
xb2=yb./(xb./(ones(gb,1)-xb)).^a;
% Berechnen des 2. Inputs der Gruppe B,
% gegeben Output und Inputmix
xb1=(xb./(ones(gb,1)-xb)).*xb2;
% Berechnen des 1. Inputs der Gruppe B,
% gegeben Output und Input 2

%%%%%% Berechnen der ineffizienten Inputs
xali=zeros(ga,1);
% Vektor definieren bzw. reseten
xa2i=zeros(ga,1);
% Vektor definieren bzw. reseten
xb1i=zeros(gb,1);
% Vektor definieren bzw. reseten
xb2i=zeros(gb,1);
% Vektor definieren bzw. reseten

xali=xali./ineffa;
% Ineffiziente Input 1 der Gruppe A
xa2i=xa2i./ineffa;
% Ineffiziente Input 2 der Gruppe A
xb1i=xb1i./ineffb;
% Ineffiziente Input 1 der Gruppe B
xb2i=xb2i./ineffb;
% Ineffiziente Input 2 der Gruppe B

X=zeros(n,2);
% Inputmatrix definieren und reseten
Y=zeros(n,1);
% Outputmatrix definieren und reseten
X=[xali, xa2i; xb1i,xb2i];
% Zusammenfassen der Inputvektoren
Y=[ya; yb];
% Zusammenfassen der Outputvektoren

%%%%%% Data Envelopment Analysis
deaineffa=zeros(ga,1);
% Vektor definieren bzw. reseten
deaineffb=zeros(gb,1);
% Vektor definieren bzw. reseten
deasel(X,Y,'I','CRS','D');
% Aufruf der DEA-Syntax, DEA-Syntax wird auf Wunsch
% der Autoren nicht publiziert

deaineffa=ans(1:ga);
% Vektor der optimalen Lösungen der Gruppe A
deaineffb=ans((ga+1):(ga+gb));
% Vektor der optimalen Lösungen der Gruppe B

%%%%%% Hypothesentests
ma=zeros(1,1);
% Skalar definieren bzw. reseten
mb=zeros(1,1);
% Skalar definieren bzw. reseten
ma=sum(deaineffa)/ga;
% Berechnen der durchschnittlichen DEA-Effizienz der Gruppe A
mb=sum(deaineffb)/gb;
% Berechnen der durchschnittlichen DEA-Effizienz der Gruppe B

[h,significance,ci]=ttest2(deaineffa, deaineffb,0.05,0);
% zweiseitiger T-Test der Resultate der DEA
dea_ht_p= [h, significance, ci];
[h,significance,ci]=ttest2(deaineffa, deaineffb,0.05,-1);
% einseitiger T-Test der Resultate der DEA
dea_ht_p1= [h, significance, ci];
[p,h,stats] = ranksum(deaineffa, deaineffb,0.05);
% Mann-Whitney-U-Test der Resultate der DEA
dea_ht_np= [p, h];

%%%%%% Sensitivitätstests
%%%%%%%% Delta Psi<1
X=zeros(n,2);
% Inputmatrix reseten
Y=zeros(n,1);
% Outputmatrix reseten
X=[xali*psiu, xa2i*psiu; xb1i, xb2i];
% Veränderung der Inputdaten der Gruppe A,
% Inputmatrix zusammenstellen
% Outputmatrix zusammenstellen
Y=[ya; yb];
% Aufruf der DEA-Syntax mit Psi<1
deasel(X,Y,'I','CRS','D');
% Berechnen der durchschnittlichen Effizienz Gruppe A Psi<1
ineffau=ans(1:ga);
% Berechnen der durchschnittlichen Effizienz Gruppe B Psi<1
ineffbu=ans((ga+1):(ga+gb));
% Berechnen der durchschnittlichen Effizienz Gruppe A Psi<1
mau=sum(ineffau)/ga;
% Berechnen der durchschnittlichen Effizienz Gruppe B Psi<1
mbu=sum(ineffbu)/gb;
% Berechnen des Validitätsmasses der Gruppe A für Psi<1
deltaua=(1-mau/ma)/(1-1/psiu);

```

```

deltaub=(mb/mbu-1)/(1-1/psiu);
deltau=abs(deltaua)+abs(deltaub);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
X=zeros(n,2);
Y=zeros(n,1);
X=[xa1i*psid,xa2i*psid; xb1i, xb2i];

Y=[ya; yb];
deasel(X,Y,'I','CRS','D');
ineffad=ans(1:ga);
ineffbd=ans((ga+1):(ga+gb));
mad=sum(ineffad)/ga;
mbd=sum(ineffbd)/gb;
deltada=(1-mad/ma)/(1-1/psid);
deltadb=(mb/mbd-1)/(1-1/psid);
deltad=abs(deltada)+abs(deltadb);

du=[deltaua, deltaub, deltau];
dd=[deltada, deltadb, deltad];
d=(deltau+deltad)/2;
X=zeros(n,2);
Y=zeros(n,1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
time1=cputime;
runtime=time1-time0;
res=zeros(1,73);
res=[r, s, q, v,g,ma, mb, ia, ib, d, du, dd, ea, eb,
      psia, psib, psiu, psid, dea_ht_p, iv_ht_p,dea_ht_np ,
      iv_ht_np, testa, testb, testa_p, testa_n,
      testb_p, testb_n, exa, exb, sxa, sxb, eya,
      eyb, sya, syb, ga, gb, runtime, dea_ht_p1, iv_ht_p1];
output=[output; res];
Iteration= [g,v,q,s,r]
Iteration= [varg,varv,varq,vars,varr]
end;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
output=output(2:(r+1),1:73);
output1=[output1; output];
output=zeros(1,73);
end;
output1=output1(2:s*r+1,1:73);
output2=[output2; output1];
output1=zeros(1,73);
end;
output2=output2(2:s*r*q+1,1:73);
output3=[output3; output2];
output2=zeros(1,73);
end;
output3=output3(2:s*r*q*v+1,1:73);
output4=[output4; output3];
output3=zeros(1,73);
end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
output4=output4(2:(s*r*q*v*g+1),1:73);
wklwrite('SIMULATION',output1);
timeall1=cputime;
output4=zeros(1,73);
Rt=timeall1-timall0
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Ende der Simulation

% Berechnen des Validitätsmasses der Gruppe B für Psi<1
% Berechnen des Validitätsmasses für Psi<1

% Inputmatrix reseten
% Outputmatrix reseten
% Veränderung der Inputdaten der Gruppe A,
% Inputmatrix zusammenstellen
% Outputmatrix zusammenstellen
% Aufruf der DEA-Syntax mit Psi>1
% Berechnen der durchschnittlichen Effizienz Gruppe A Psi>1
% Berechnen der durchschnittlichen Effizienz Gruppe B Psi>1
% Berechnen der durchschnittlichen Effizienz Gruppe A Psi>1
% Berechnen der durchschnittlichen Effizienz Gruppe B Psi>1
% Berechnen des Validitätsmasses der Gruppe A für Psi>1
% Berechnen des Validitätsmasses der Gruppe B Psi>1
% Berechnen des Validitätsmasses für Psi>1

% Speichern der Validitätsmasse für Psi<1
% Speichern der Validitätsmasse für Psi>1
% Berechnen des Mittelwertes der Validitätsmasse
% Inputmatrix reseten
% Outputmatrix reseten

% Endzeit der Iteration speichern
% Berechnen der Rechenzeit der Iteration
% Vektor definieren und reseten
% Speichern der Daten und Resultate der Iteration

% Zusammenfügen der Vektoren der Resultate

% Datei exportieren
% Endzeit der Simulation
% Matrix löschen
% Berechnen der Simulationszeit

```

Anhang B

Fehlerwahrscheinlichkeiten

Tabelle B.1: Fehlerwahrscheinlichkeiten, 95%-Signifikanzniveau ($N_1 = N_2 = 100$)

| $\iota - \mathcal{B}^{AB}$ | Fehler erster Art | | |
|----------------------------|-------------------|-------------------|-------------------|
| | $\sigma_i = 0.15$ | $\sigma_i = 0.20$ | $\sigma_i = 0.25$ |
| 0.00 | 0.0500 | 0.0500 | 0.0500 |

| $\iota - \mathcal{B}^{AB}$ | Fehler zweiter Art | | |
|----------------------------|--------------------|-------------------|-------------------|
| | $\sigma_i = 0.15$ | $\sigma_i = 0.20$ | $\sigma_i = 0.25$ |
| 0.00 | 0.9750 | 0.9750 | 0.9750 |
| 1.00 | 0.9317 | 0.9459 | 0.9532 |
| 2.00 | 0.8455 | 0.8949 | 0.9184 |
| 3.00 | 0.7074 | 0.8158 | 0.8668 |
| 4.00 | 0.5296 | 0.7074 | 0.7963 |
| 5.00 | 0.3457 | 0.5762 | 0.7074 |
| 6.00 | 0.1926 | 0.4359 | 0.6037 |
| 7.00 | 0.0902 | 0.3033 | 0.4921 |
| 8.00 | 0.0351 | 0.1926 | 0.3810 |
| 9.00 | 0.0112 | 0.1109 | 0.2791 |
| 10.00 | 0.0029 | 0.0576 | 0.1926 |
| 11.00 | 0.0006 | 0.0269 | 0.1248 |
| 12.00 | 0.0001 | 0.0112 | 0.0758 |
| 13.00 | 0.0000 | 0.0042 | 0.0430 |
| 14.00 | 0.0000 | 0.0014 | 0.0228 |
| 15.00 | 0.0000 | 0.0004 | 0.0112 |
| 16.00 | 0.0000 | 0.0001 | 0.0052 |
| 17.00 | 0.0000 | 0.0000 | 0.0022 |
| 18.00 | 0.0000 | 0.0000 | 0.0009 |
| 19.00 | 0.0000 | 0.0000 | 0.0003 |
| 20.00 | 0.0000 | 0.0000 | 0.0001 |

Literaturverzeichnis

- Banker, R. D. (1993). Maximum Likelihood, Consistency and Data Envelopment Analysis. *Management Science*, 39(10):1265–73.
- (1996). Hypothesis Tests Using Data Envelopment Analysis. *Journal of Productivity Analysis*, 7(2-3):139–59.
- Banker, R. D., A. Charnes und W. W. Cooper (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30:1078–92.
- Banker, R. D., R. F. Conrad und R. P. Strauss (1986). A Comparative Application of DEA and Translog Methods: An Illustrative Study of Hospital Production. *Management Science*, 32:30–44.
- Banker, R. D. und R. C. Morey (1986). Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operations Research*, 34(4):513–21.
- Burgess, J. F. und P. W. Wilson (1996). Hospital Ownership and Technical Inefficiency. *Management Science*, 42:110–23.
- Byrnes, P. und V. Valdmanis (1994). Analyzing Technical and Allocative Efficiency of Hospitals. In A. Charnes und et al., Hg., *Data Envelopment Analysis: Theory, Methodology, and Application*, 129–44. Boston, Dordrecht, London: Kluwer Academic Publisher.
- Charnes, A., W. W. Cooper, Z. M. Huang und D. B. Sun (1990). Polyhedral Cone-Ratio DEA Models with an Illustrative Application to Large Commercial Banks. *Journal of Econometrics*, 46(1-2):73–91.

- Charnes, A., W. W. Cooper und E. Rhodes (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2(6):429–44.
- (1981). Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through. *Management Science*, 27(6):668–97.
- Chiang, A. C. (1984). *Fundamental Methods of Mathematical Economics*. New York: McGraw-Hill, 3 Auflage.
- Coelli, T., D. S. P. Rao und G. E. Battese (1998). *An Introduction to Efficiency and Productivity Analysis*. Boston, Dordrecht, London: Kluwer Academic.
- Cooper, W. W., K. S. Park und J. T. Pastor (1999). RAM: A Range Adjusted Measure of Inefficiency for Use with Additive Models, and Relations to Other Models and Measures in DEA. *Journal of Productivity Analysis*, 11(1):5–42.
- Cooper, W. W., L. M. Seiford und K. Tone (2000). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Boston, Dordrecht, London: Kluwer Academic Publishers.
- Crivelli, L. (1998). *Grenzüberschreitende Patientenwanderungen*. Bern, Stuttgart, Wien: Haupt.
- Danzig, G. B. (1951). Maximization of a Linear Function of Variables Subject to Linear Inequalities. In T. C. Koopmans, Hg., *Activity Analysis of Production and Allocation*, 339–47. New York, London: John-Wiley and Sons, Inc.
- Debreu, G. (1951). The Coefficient of Resource Utilization. *Econometrica*, 19(3):273–92.
- Deprins, D., L. Simar und H. Tulkens (1984). Measuring Labor Efficiency in Post Offices. In M. Marchand, P. Pestieau und H. Tulkens, Hg., *The Performance of Public Enterprises: Concepts and Measurement*, 243–67. Amsterdam: North Holland.
- Färe, R., S. Grosskopf, B. Lindgren und P. Roos (1994). Productivity Developments in Swedish Hospitals: A Malmquist Output Index Approach. In A. Charnes, W. W. Cooper, A. Y. Lewin und L. M. Seiford, Hg., *Data Envelopment Analysis. Theory, Methodology and Applications*, 253–72. Boston, Dordrecht, London: Kluwer Academic Publishers.

- Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*, 120:253–81.
- Fried, H. O., C. A. K. Lovell und P. Vanden Eeckaut (1993). Evaluating the Performance of U.S. Credit Unions. *Journal of Banking and Finance*, 17:251–65.
- Greene, W. H. (1997). *Econometric Analysis*. New Jersey: Prentice Hall.
- Grosskopf, S. (1996). Statistical Inference and Nonparametric Efficiency: A Selective Survey. *Journal of Productivity Analysis*, 7(2-3):161–76.
- Kennedy, P. (1998). *A Guide to Econometrics*. Cambridge, Massachusetts: The MIT Press, 4 Auflage.
- Kittelsen, S. A. C. (1999). Monte Carlo Simulations of DEA Efficiency Measures and Hypothesis Testing. Memorandum 09/99, Departement of Economics, University of Oslo.
- Kneip, A., U. Park und L. Simar (1998). A Note on the Convergenze of Nonparametric DEA-Estimators for Production Efficiency Scores. *Econometric Theory*, 14:783–93.
- Koopmans, T. C. (1951). An Analysis of Production as an Efficient Combination of Activities. In T. C. Koopmans, Hg., *Activity Analysis of Production and Allocation*, 33–97. New York, London: John-Wiley and Sons, Inc.
- Kumbhakar, S. C. und C. A. K. Lovell (2000). *Stochastic Frontier Analysis*. New York, Oxford, Toronto, Melbourne: Cambridge University Press.
- Laffont, J. J. und J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation*. Cambridge and London: MIT Press.
- Linna, M. und U. Hakkinen (1998). A Comparative Application of Econometric Frontier and DEA Methods for Assessing Cost Efficiency of Finnish Hospitals. In P. Zweifel, Hg., *Health, the Medical Profession, and Regulation. Developments in Health Economics and Public Policy, vol. 6*, 169–87. Boston, Dordrecht, London: Kluwer Academic.
- Lovell, C. A. K. (1993). Production Frontiers and Productive Efficiency. In H. O. Fried, C. A. K. Lovell und S. S. Schmidt, Hg., *The Measurement of Productive Efficiency*:

- Techniques and Applications*, 3–67. New York, Oxford, Toronto, Melbourne: Oxford University Press.
- Marsaglia, G. und W. W. Tsang (1984). A Fast, Easily Implemented Method for Sampling from Decreasing or Symmetric Unimodal Density Functions. *SIAM Journal of Scientific and Statistical Computing*, 5:349–59.
- Mas-Colell, A., M. D. Whinston und J. R. Green (1995). *Microeconomic Theory*. New York, Oxford, Toronto, Melbourne: Oxford University Press.
- McCarty, T. A. und S. Yaisawarng (1993). Technical Efficiency in New Jersey School Districts. In H. O. Fried, C. A. K. Lovell und S. S. Schmidt, Hg., *The Measurement of Productive Efficiency: Techniques and Applications*, 271–87. New York, Oxford, Toronto, Melbourne: Oxford University Press.
- OECD (1999). *OECD Health Data*. Paris.
- Olesen, O. B. und N. C. Petersen (1991). Collinearity in Data Envelopment Analysis: An Extended Facet Approach. Workin paper no. 1/1991, Departement of Management, University of Odense.
- Pareto, V. (1897). *Cours d'économie politique*, Band 2. Lausanne.
- Ray, S. C. (1988). Data Envelopment Analysis, Nondiscretionary Inputs and Efficiency: An Alternative Interpretation. *Socio-Economic Planning Sciences*, 22(4):167–76.
- Schnell, R., P. Hill und E. Esser (1999). *Methoden der empirischen Sozialforschung*. München, Wien, Oldenbourg: Oldenbourg.
- Simar, L. und P. W. Wilson (2000). Statistical Inference in Nonparametric Frontier Models: The State of the Art. *Journal of Productivity Analysis*, 13(1):49–78.
- Stata (2001). *Stata Statistical Software: Release 7.0*. College Station, TX.
- Steinmann, L. und P. Zweifel (2000). Zur (In)Effizienz schweizerischer Krankenhäuser. (On the [In]Efficiency of Swiss Hospitals. With English summary.). *Ifo Studien*, 46(2):197–217.
- (2001). The Range Adjusted Measure (RAM) in DEA: Comment. *Journal of Productivity Analysis*, 15(2):139–44.

- Thompson, R. G., L. Langemeier, C. T. Lee, E. Lee, und R. M. Thrall (1990). The Role of Multiplier Bounds in Efficiency Analysis with Application to Kansas Farming. *Journal of Econometrics*, 46(1-2):93–108.
- Timmer, C. P. (1971). Using a Probabilistic Frontier Production Function to Measure Technical Efficiency. *Journal of Political Economy*, 79(4):776–94.
- Varian, H. R. (1995). *Microeconomic Analysis*. New York: Norton, 3 Auflage.
- Zweifel, P. und F. Breyer (1997). *Health Economics*. New York: Oxford University Press.

Lebenslauf

| | |
|---------------|---|
| Name: | Steinmann |
| Vorname: | Lukas Adrian |
| Geburtsdatum: | 23. August 1971 |
| Geburtsort: | Zürich |
| Bürgerort: | Zürich |
| Zivilstand: | verheiratet mit Hanna Elisabeth Andersdotter, geborene Wulf aus Vänersborg, Schweden |
| Ausbildung: | Primarschule in Bethesda, Maryland, USA, 1976-77 Primarschule in Witikon ZH, 1977-84 Sekundarschule in Witikon ZH, 1984-86 Gymnasium Typus E in Hottingen ZH, 1987-91 Studium an der Universität Zürich, 1992-98 Lizentiat in Volkswirtschaftslehre, 1999 Studium an der Universität Zürich, 1999-2002 Doktorat in Volkswirtschaftslehre, 2002 |